

VÝVOJ TEMATICKEJ PRESNOSTI VRSTVY HUSTOTY NEPRIEPUSTNOSTI PRE ROKY 2006 A 2009 NA SLOVENSKU A V ČESKU

Andrea Čorbová^{*,**}, Pavol Hurbánek^{*}, Konštantín Rosina^{*}

^{*} Geografický ústav SAV, v. v. i., Štefánikova 49, 814 73 Bratislava, Slovensko

^{**} Univerzita Komenského v Bratislave, Prírodovedecká fakulta, Ilkovičova 6, 842 15 Bratislava, Slovensko,
and.corbova@gmail.com, pavolhurbánek@gmail.com, konstantin.rosina@savba.sk

Thematic accuracy development of Imperviousness Density Layer for years 2006 and 2009 in Slovakia and Czechia

The Imperviousness Density Layer (IMD) of the Copernicus Programme is a raster with 100 m pixel resolution and pixel integer values from 0 to 100 % representing the imperviousness degree. So far, the IMD validations have been based on stratified systematic samples with relatively small sample sizes, in the case of the joint area of Slovakia and Czechia totalling about several hundred sample pixels, each with 25 sampling points. The objective of this paper is to evaluate the IMD thematic accuracy based on a large simple random sample of 20,000 pixels per country per year, each pixel with 100 sampling points. A large sample is capable of a more precise estimate of the omission error compared to the smaller samples used before (when dealing with a small proportion class), which is important e.g. in population disaggregation (dasymetric mapping). The focus is on the IMD for the reference years 2006 and 2009 in Slovakia and Czechia and how this accuracy developed with each new version. The analysis reveals gradual improvement in most (but not all) of the different aspects of the IMD thematic accuracy with almost each new version, especially with the latest ones (published in 2019) compared to the previous ones (published in 2009, 2010 and 2013). However, the estimates of the proportions of impervious surfaces from the total area of Slovakia and Czechia based on these latest versions seem to be the worst compared to the previous ones. Also, the area of impervious surfaces incorrectly classified as pervious in a large number of pixels with zero imperviousness map values and small nonzero imperviousness reference values (i.e. in 'major underestimation error pixels' or 'omission error pixels with a threshold value set to 1%') remains large, and its proportion of the total area of incorrectly classified surfaces reaches its maximum in the latest versions.

Key words: IMD, imperviousness, validation, thematic accuracy, Slovakia, Czechia

ÚVOD

V mnohých oblastiach ľudskej činnosti zaujali veľmi dôležité postavenie čo najpresnejšie informácie o priestorovej diferenciacii rozmiestnenia obyvateľstva. Vedomosť o rozmiestnení obyvateľstva je veľmi dôležitá najmä pri plánovaní využitia krajiny, napr. pri lokalizácii nových bytových jednotiek, nemocníc, nákupných stredísk atď., ale aj pri procesoch tvorby krízových scenárov či pri procesoch zmierňovania následkov prírodných katastrof (GSE Land 2008).

Veľmi dobrým indikátorom prítomnosti a priestorového rozmiestnenia ľudí sú práve nepriepustné povrchy, na ktorých sa obyvateľstvo sústreďuje. Údaje o výskyte nepriepustných povrchov, a teda aj údaje o predpokladanej prítomnosti ľudí, sú v súčasnosti najčastejšie získavané z dát z diaľkového prieskumu Zeme (DPZ). Tieto dáta sa zbierajú v pravidelných časových intervaloch, a preto je možné na sledovaných územiach nielen pozorovať zaznamenaný stav, ale aj porovnávať zmeny, ktoré tam nastali (EEA 2006).

Vrstva nepriepustnosti s vysokým rozlíšením (Imperviousness High Resolution Layer – IHRL), skrátene tiež vrstva hustoty nepriepustnosti (Imperviousness Density – IMD), je jednou z celoeurópskych služieb programu Copernicus (Copernicus Land Monitoring Service – CLMS), resp. predtým programu Globálne monitorovanie pre životné prostredie a bezpečnosť (Global Monitoring for Environment and Security – GMES) – EEA (2020). Je to rastrová vrstva so 100-metrovým agregovaným validovaným a 20-metrovým (resp. 10-metrovým v prípade IMD2018) plným nevalidovaným rozlíšením, v ktorej každý pixel nadobúda celočíselnú hodnotu z množiny (0, 1, 2, 3, ..., 100) reprezentujúcu percentuálny podiel nepriepustných povrchov. Hustota nepriepustnosti je vypočítaná automatizovane pomocou algoritmu založeného na kalibrovanom normalizovanom diferenčnom vegetačnom indexe – NDVI (Smith 2020).

Tri hlavné kroky spracovania dát sú popísané v správe o doručení (Kopecký a Kahabka 2008, p. 8):

- hybridná automatická klasifikácia, ktorej výsledkom bola binárna mapa zastavaných areálov;
- manuálna korekcia binárnej mapy zastavaných areálov tak, aby bola dosiahnutá 85-percentná tematická presnosť, rovnako ako dobré kvalitatívne výsledky;
- odvodenie stupňa nepriepustnosti na základe NDVI (ukazovateľ popisujúci prítomnosť vegetácie na základe odrazivosti v červenom a blízkom infračervenom spektrálnom pásme).

Podľa definície zo systému klasifikácie krajiny pokrývky (FAO 2005) „nepriepustné plochy sú charakterizované nahradením pôvodnej prírodnej alebo poloprirodnej pokrývky či vodnej plochy umelým, často nepriepustným povrchom; umelosť je obvykle charakteristická dlhou trvácnosťou“ (EEA 2007, p. 2). Táto vrstva pokrýva 39 krajín Európskeho hospodárskeho priestoru a bola prvýkrát vytvorená pre rok 2006 (vtedy pokrývala 38 krajín) popri aktualizácii už desaťročia známeho projektu CORINE Land Cover (CLC), na základe rovnakých vstupných dát z DPŽ ako pre CLC triedu umelých povrchov. IMD vznikla ako riešenie problému s minimálnym identifikovaným areálom 25 ha, ktorý má CLC. Od roku 2006 sa Európska environmentálna agentúra (EEA) v spolupráci s Európskou vesmírnou agentúrou (ESA) a Európskou komisiou (EK) snažia o jej generovanie a aktualizáciu v pravidelných trojročných intervaloch (EEA 2020). V súčasnosti je dostupná pre roky 2006, 2009, 2012, 2015 a 2018 (Smith 2020).

Pri vývoji IMD sa však postupne ukázalo, že nie je ľahké naplniť to, čo si európske inštitúcie stanovili za cieľ pri využívaní IMD, a teda, že sa vznikom tejto novej rastrovej vrstvy malo začať s presnejšími analýzami a riešením rôznych geografických, ekonomických či sociologických problémov (EK 2015), pri ktorých by tento produkt mohol zohrávať dôležitú úlohu. Národné referenčné centrá (NRC) viacerých krajín sa k prvým verziám stavali rezervovane odvolávajúc sa pritom na ich problematickú tematickú presnosť. Slovenské NRC Slovenská agentúra životného prostredia (SAŽP), ktorá spolupracuje s EEA, vykonala kvalitatívne hodnotenie jednej z prvých verzií tohto produktu s takýmto výsledným komentárom: „zatiaľ na GIS aplikácie nepoužiteľný a málo vierohodný ... vyskytujú sa v ňom hrubé nedostatky oproti realite a SAŽP ho neodporúčala použiť v EEA na akékoľvek analýzy, ktoré sú dôležité pre rozhodovacie procesy Európskej komisie a majú legislatívne, reštriktívne alebo finančné dosahy na členské štáty“ (Machková a Nováček 2009, p. 13).

Je zjavné, že postupný vývoj jednotlivých aktualizáčnych verzii IMD vedie ku postupnému zlepšovaniu ich tematickej presnosti, no miera a rozličné aspekty tohto zlepšovania neboli doposiaľ známe. Cieľom príspevku je preto analyzovať vývoj rôznych aspektov tematickej presnosti jednotlivých doteraz zverejnených verzii vrstvy hustoty nepriepustnosti IMD pre roky 2006 a 2009 na Slovensku a v Česku s dôrazom na ich využitie pri priestorovej dezagregácii sociálno-ekonomických, najčastejšie populačných premenných, napríklad ako pomocné dáta pri dazymetrickom mapovaní. Práve s týmto cieľom boli niektoré aktualizáčne verzie IMD2006 a IMD2009 v minulosti využité, napr. Rakúskym technologickým inštitútom (Steinnocher et al. 2011). Spoločným výskumným centrom Európskej komisie (Batista e Silva et al. 2013) a ďalšími (napr. Rosina a Hurbánek 2016). Je preto žiaduce, aby odborná obec používajúca populačné gridy pre dané roky ako výstupy týchto prác (príp. ďalších, ktoré vznikli alebo ešte len budú vznikať) poznala kvalitu príslušných aktualizáčnych verzii IMD2006 a IMD2009 ako vstupov do týchto prác. Na rozdiel od IMD z neskoršieho obdobia (pre roky 2012, 2015 a 2018), ktoré boli vďaka výraznému nárastu dostupnosti alternatívnych dátových produktov o rozmiestnení nepriepustných povrchov v druhej dekáde tohto milénia viackrát validované (napr. Congedo et al. 2016, Smith 2020 a Lehner et al. 2017), IMD pre roky 2006 a 2009 v tomto ohľade zaostávajú. Hlavnou príčinou je zrejme nedostatok alternatívnych dátových produktov pre tieto dva roky (ktoré by mohli byť použité ako referenčné), a teda náročnosť takejto validácie vyžadujúcej vytvorenie vlastnej referenčnej databázy na tento konkrétny účel.

VÝVOJ VRSTVY HUSTOTY NEPREPUSTNOSTI (IMD)

Základom úsilia Európskej únie o monitorovanie Zeme a jej ekosystémov je program Copernicus. Zaisťuje sa ním, že jej občania budú pripravení na krízy a prírodné a človekom spôsobené katastrofy a chránení pred nimi. V rámci služieb programu Copernicus sa podporuje široká škála environmentálnych a bezpečnostných aplikácií vrátane monitorovania zmeny klímy, trvalo udržateľného rozvoja, dopravy a mobility, regionálneho aj miestneho plánovania, námorného dozoru, poľnohospodárstva a zdravia. Služby programu Copernicus sú založené na informáciách zo špecializovanej sústavy družíc známej ako „Sentinel“, ako aj z desiatok družíc tretích strán označených ako „prispievajúce vesmírne misie“, ktoré dopĺňajú údaje z meraní „in situ“ (teda miestne alebo na mieste). Službami sa transformuje toto veľké množstvo údajov na informácie s pridanou hodnotou, a to spracovaním a analýzou údajov, ich integráciou do iných zdrojov a potvrdením výsledkov. Dátové súbory sa za celé roky a desaťročia upravujú tak, aby boli porovnateľné a dalo sa v nich vyhľadávať, čím sa zaisťuje monitorovanie zmien; skúmajú sa modely, ktoré sa používajú na vytvorenie lepších prognóz, napríklad v súvislosti s oceánom a atmosférou. Zo snímok sa vytvárajú mapy, identifikujú sa znaky a anomálie a získavajú sa štatistické informácie.

Vrstva hustoty nepriepustnosti s vysokým rozlíšením IMD2006 bola prvou, ktorá bola vygenerovaná v rokoch 2006 – 2008 z multisenzorových, bitemporálnych a ortorektifikovaných satelitných snímok, a to popri aktualizácii CORINE Land Cover 2006. Jej generovanie sa týkalo 38 európskych krajín (32 členských štátov EHP a 6 krajín západného Balkánu). Na začiatku roku 2007 bol vypísaný tender EEA/IDS/07/001 na zhotovenie prvej verzie produktu IMD v hodnote 1,2 – 2 mil. eur. Predmetom kontraktu bola „produkcia konzistentnej celoeurópskej databázy zastavaných plôch s vysokým rozlíšením, obsahujúcej hustotu nepriepustnosti pre

referenčný rok 2006“ na báze ortorektifikovaných satelitných snímok poskytnutých z ESA (EEA 2007, p. 2). Účelom kontraktu bola „podpora širokej škály požiadaviek na informácie o životnom prostredí pre viaceré oblasti politik na európskej, národnej a regionálnej úrovni, vrátane poľnohospodárstva, kvality pôd, vody, ovzdušia a mestského prostredia, ochrany prírody...“ (EEA 2007, p. 1).

Od začiatku produkcie sa vytvoril časový rad stavových vrstiev hustoty nepriepustností v celej oblasti pokrytej 39 členskými a spolupracujúcimi krajinami Európskej environmentálnej agentúry (EEA), ktoré predstavujú celkovo 6 miliónov km² (EEA 2020).

Vrstva pre prvé dva referenčné roky (IMD2006 a IMD2009) bola pôvodne vyrobená v rámci aktivít európskeho projektu Geoland2 a GMES. Počas generovania IMD2012, keď sa začala súčasne generovať aj prvá zmenová vrstva (vrstva zachytávajúca len prírastok alebo úbytok nepriepustných povrchov medzi dvoma referenčnými rokmi) medzi rokmi 2009 a 2012, sa zistilo, že je veľmi pravdepodobné, že došlo k určitému nadhodnoteniu hodnôt podielu hustoty nepriepustných povrchov v IMD2006 a IMD2009. Prišlo sa totiž na to, že absolútne hodnoty nepriepustnosti v roku 2009 boli v mnohých oblastiach vyššie ako v roku 2012, hoci zmenová vrstva popisovala práve zvýšenie nepriepustnosti v týchto oblastiach (EEA 2017). To bol dôvod, prečo služba Copernicus následne prepracovávala postupne tieto referenčné roky. V rámci produkcie vrstvy hustoty nepriepustnosti k referenčnému roku 2015 (IMD 2015) boli preto opätovne spracované a navzájom harmonizované stavové vrstvy a vrstvy zmien vzťahujúce sa na všetky predchádzajúce referenčné roky 2006, 2009 a 2012, a teda vznikli novšie, vylepšené verzie pre tieto referenčné roky (EEA 2017).

Opisu metodológie tvorby aktualizovanej verzie referenčného roka 2009 sa venovali Gangkofner et al. (2010) a aktualizácii referenčného roka 2012 Lefebvre et al. (2013). Podobné metódy boli aplikované aj v USA (2011) na vývoj databázy National Land Cover (Xian et al. 2011).

Najnovšie špecifikácie opravených verzií vrstvy hustoty nepriepustnosti ponúkajú Langanke et al. (2016), ktorí popisujú zlepšenia berúc do úvahy skúsenosti z predchádzajúcich výsledkov overenia produktov týkajúce sa zabezpečenia správnej klasifikácie priepustných a nepriepustných povrchov.

Pri úprave vrstvy hustoty nepriepustnosti bola navyše použitá aj metóda vyplňania medzier (EEA 2017), ktorou sa malo zabrániť medzerám vo vrstvách spôsobených chýbajúcimi vstupnými údajmi pri pozorovaní Zeme, napr. prítomnosťou oblakov nad daným územím v čase snímkovania. Metóda spočíva v tom, že medzery, ktoré boli zachytené vo vrstve z referenčného roku 2009, sa doplnili údajmi z referenčného roku 2006 a medzery z referenčného roku 2012 sa vyplnili údajmi z referenčného roku 2009 atď. Predpokladá sa teda, že práve v týchto oblastiach môže dôjsť k určitému podhodnoteniu nárastu nepriepustných povrchov v danom referenčnom roku, keďže sa využívajú údaje z predchádzajúcich referenčných rokov.

V tomto prípade je vhodné presnejšie definovať pojem referenčný rok/obdobie podľa EEA (2006, p. 4): „za referenčný rok/obdobie sa pokladá časové obdobie najviac tri roky“ (rok pred referenčným rokom, referenčný rok a rok po referenčnom roku). To znamená, že údaje pozorovania Zeme použité na vytvorenie vrstvy IMD2009 sú v skutočnosti kombináciou snímkovania z rokov 2008, 2009 a 2010.

Údaje o nepriepustnosti sú dostupné pre referenčné roky 2006, 2009, 2012, 2015 a 2018 a obsahujú dva typy produktov – stavové vrstvy a vrstvy zmien

(PROJECT MANAGER COPERNICUS LAND SERVICES 2020). Tento príspevok je venovaný len stavovým vrstvám pre roky 2006 a 2009 na Slovensku a v Česku.

PROBLEMATIKA POJMU NEPRIEPUSTNOSTI A VÝZNAM POTREBY VALIDÁCIE

IMD bola od svojho vzniku prísľubom pozitívneho vývoja do budúcnosti. Mapovanie rozmiestnenia nepriepustných povrchov je dôležité hlavne pre ochranu životného prostredia, ochranu pôd, poľnohospodársku politiku a pre fyzickú geografiu (EEA 2006, p.4). Využitie tejto vrstvy možno však nájsť aj v humánnej geografii. Tak ako CLC aj IMD môže byť využitá ako pomocná vrstva pre dezagregáciu prakticky ľubovoľných sociálno-ekonomických premenných vrátane populačného modelovania (Rosina a Hurbánek 2016). Avšak, je potrebné, aby bolo známe, aká je tematická presnosť jednotlivých verzií IMD.

IMD bola pôvodne na začiatku tvorby označovaná ako Európska mapa hustoty nepriepustnosti, celým názvom Fast Track Service Precursor on Land Monitoring – Degree of Soil Sealing. Pri vypísaní prvého tendra v roku 2006, v ktorom boli špecifikované požiadavky na generovanie IMD, sa uvádza, že na účely validácie sa za zastavaný areál bude považovať územie s podielom nepriepustných povrchov (degree of soil sealing) 80 – 100 %, čím sa zavádza tzv. prah (threshold), na základe ktorého sa pomerové (metrické) hodnoty 0 – 100 % transformujú na dichotomické hodnoty: zastavané (built-up) a nezastavané (non-built-up) – Hurbánek et al. (2010). Maucha et al. (2010) neskôr navrhujú zmeniť tento prah z 80 % na 30 %. Odôvodňujú to tým, že validácia s prahom 30 %, na rozdiel od validácie s pôvodne navrhovaným prahom 80 %, vedie k chybám vynechania a priradenia menším ako 15 %, čo bola požiadavka definovaná v tendri.

Podstatou validácie je overenie, či určitý produkt spĺňa stanovené kritériá. Na rozdiel od verifikácie by sa mala validácia vykonávať až po ukončení procesu produkcie a viesť k prijatiu alebo zamietnutiu databázy a nie k jej modifikácii. V tomto prípade bola validácia chápaná ako overenie, či tematická mapa odvodená z IMD spĺňa kritérium 85 % tematickej presnosti stanovenej v podmienkach tendra na zhotovene produktu. Maucha a Büttner (2008) konštatujú, že iba skontrolovaním všetkých pixelov mapy a zistením, že presne 15 % z nich bolo nesprávne klasifikovaných, by bolo možné preukázať, že tematická presnosť mapy je presne 85 %. Avšak skontrolovať každý pixel by bolo finančne a časovo nákladnejšie ako samotná produkcia hodnotenej mapy, preto je nevyhnutné použiť na účel validácie štatistické metódy výberu vzorky. Štatistický výber ako typickú metódu hodnotenia máp krajinej pokrývky uvádza aj Stehman (1999, p. 2425): „Hodnotenie presnosti je typicky založené na výbere referenčných lokalít, v ktorých je skutočná (referenčná) klasifikácia krajinej pokrývky porovnaná s klasifikáciou v hodnotenej mape“.

Posúdenie tematickej presnosti je dôležité aj z pohľadu užívateľov tejto vrstvy, ktorí ju využívajú pri aplikácii v rôznych oblastiach pri riešení rozličných problémov, aby vedeli, do akej miery je správna, resp. spoľahlivá.

Podľa Kolára (2003) kvalita hodnôt v databáze je spravidla posudzovaná na základe štatistického porovnania jednotlivých hodnôt s referenčnými údajmi, ktoré majú vyššiu kvalitu a pochádzajú z nezávislého zdroja. Zistiť tematickú presnosť skúmaného objektu podľa Kolára (2003) znamená uviesť pravdepodobnosť, s akou

má objekt skutočne vlastnosť, ktorá je udávaná príslušným atribútom. Základné požiadavky na uskutočnenie validácie sú jasne definované kritériá.

Jednou z úloh NRC, v prípade Slovenska (SAŽP), bolo po doručení prvej verzie IMD2006 spracovať kvalitatívne hodnotenie (Steenmans a Sousa 2007). Kvalita dát bola posudzovaná v rámci 18 vybraných tried CLC, hodnotila sa na päťstupňovej verbálnej stupnici – veľmi zlá, nedostatočná, prijateľná, dobrá, excelentná. Z 18 tried bolo na Slovensku posudzovaných 13 (zvyšných päť sa nevyskytuje). Iba v jednej z tried, aj to len málo významnej (prístavy), bolo hodnotenie „dobrý“, pri štyroch druhoch pokrývky (krovinné a bylinné vegetačné spoločenstvá, plochy s ojedinelou vegetáciou, holé skaly a vnútrozemské močiare) bolo hodnotenie „prijateľný“ a v ôsmich prípadoch (urbanizované štruktúry; priemyselné a obchodné jednotky; cestná a železničná sieť; letiská; ťažba nerastov, skládky a lokality vo výstavbe; orná pôda; lesy a rôznorodé poľnohospodárske plochy) bolo hodnotenie „nedostatočný“. Skutočnosťou, že presnosť vo väčšine z kontrolovaných tried bola zhodnotená ako nedostatočná, zodpovedá aj celkové kvalitatívne hodnotenie. Produkt nebol odporúčaný na použitie pre spoľahlivé modelovanie, GIS aplikácie, strategické plánovanie či reguláciu. Autori okrem iného vidia problém aj v nejasnej definícii nepriepustnosti a zastavanosti a celkovej nespoľahlivosti klasifikácie v rámci hlavných tried. Problémom môže byť prekrytie umelých objektov tenkým porastom zelených rastlín (zrejme mali na mysli machy, lišajníky a pod), takže sú identifikované ako vegetácia (Machková et al. 2008).

Dávno pred vypísaním tendra definovala EEA pojem nepriepustnosti veľmi obšírne. Pod termínom „soil sealing“ (znepriepustnenie, zapečatenie pôdy) sa rozumie zmenenie vlastností pôdy, ktorá sa tým mení na nepriepustné médium, napr. zhutnením poľnohospodárskou mechanizáciou. Zároveň sa však tento istý pojem používa na popísanie prekrytia pôdneho povrchu nepriepustnými materiálmi ako je betón, kov, sklo, asfalt, makadam (EEA 2009, p. 2). Ak by sa pod pojmom „soil sealing“ malo chápať len prekrytie pôdneho povrchu nepriepustnými materiálmi (t. j. druhá časť uvedenej definície), používatelia v reakciách na prvú verziu produktu navrhli nahradiť tento pojem pri ďalších verziách pojmom „imperviousness“ (nepriepustnosť). Používatelia vyjadrili tiež potrebu presnejšej definície toho, čo IMD mapuje. Žiaľ, ďalší tender vypísaný v roku 2009 na vylepšenie pôvodnej IMD z roku 2006 reagoval len na prvú z týchto dvoch požiadaviek a nahrádza pojem „soil sealing“ pojmom „impreviousness“ (EEA 2009, p. 2). Nejasnosť obsahu starého pojmu „soil sealing“, resp. nového „impreviousness“, teda toho, aké konkrétne plochy tento pojem označuje, odstránená nebola. Do roku 2020 sa pre tento produkt používa označenie Imperviousness Layer (IL). Zverejnením najnovšej verzie tohto produktu pre referenčný rok 2018 v roku 2020, spolu s úplne novým produktom Impervious Built-up – IBU (Smith 2020), sa začína používať aktuálne označenie Imperviousness Density (IMD).

HODNOTENIE TEMATICKEJ PRESNOSTI ZO STRANY PRODUCENTA

Súčasťou dodávky prvej verzie IMD boli aj správy o doručení pre jednotlivé štáty zhotovené spoločnosťou, ktorá spracovávala územie príslušného štátu. Okrem technickej špecifikácie dodaných dát a stručného komentáru k metodike ich tvorby obsahujú tieto správy aj hodnotenie tematickej presnosti dichotomickej odvodeniny tohto produktu aplikáciou prahu 80 % tak, ako to bolo už spomenuté. V prvej verzii IMD2006 bolo takto na Slovensku hodnotených 73 náhodne vybraných pixelov.

Vytvorená bola chybová matica (tab. 1), vypočítaná používateľská správnosť (pomer počtu správne klasifikovaných pixelov určitej kategórie s počtom pixelov danej kategórie v hodnotenej vrstve; t. j. pravdepodobnosť, že pixel klasifikovaný v mape do určitej kategórie patrí do tejto kategórie aj v skutočnosti), chyba priradenia (doplňok používateľskej správnosti do 100 %), spracovateľská správnosť (pomer počtu správne klasifikovaných pixelov určitej kategórie k počtu pixelov danej kategórie v referenčnej databáze; t. j. správnosť, s akou bola zmapovaná daná trieda), chyba vynechania (doplňok spracovateľskej správnosti do 100 %) a celková správnosť (pomer sumy správne klasifikovaných pixelov k celkovému počtu hodnotených pixelov) – Kopecký a Kahabka (2008).

Tab. 1. Výsledok validácie producentom vo forme chybovej matice

| | Hodnotená vrstva | | | Spracovateľská správnosť | Chyba vynechania | |
|-----------------|-------------------------|--------|---------|--------------------------|------------------|-------|
| | ≥ 80 % | < 80 % | Σ | | | |
| Referenčné dáta | ≥ 80 % | 2 | 0 | 2 | 100,0 % | 0,0 % |
| | < 80 % | 3 | 68 | 71 | 95,8 % | 4,2 % |
| | Σ | 5 | 68 | 73 | | |
| | Používateľská správnosť | 40,0 % | 100,0 % | | | |
| | Chyba priradenia | 60,0 % | 0,0 | | | |
| | Celková správnosť | 95,9 % | | | | |

Zdroj dát: Kopecký a Kahabka (2008).

Maucha a Büttner (2008) navrhujú použiť stratifikovaný náhodný výber s 2 000 vzorkami z pixelov IMD s hodnotami menšími ako 80 % a s 500 vzorkami z pixelov IMD s hodnotami 80 % alebo väčšími. Výsledky takéhoto hodnotenia v šiestich európskych krajinách vrátane Slovenska uvádzajú Hurbánek et al. (2010, p. 388). Ako však priznávajú Maucha et al. (2010, p. 15), dosiahnuť stratifikovaným náhodným výberom uspokojivý odhad chyby vynechania pre triedy s malým zastúpením na celkovej rozlohe záujmového územia, akou je aj trieda nepriepustných povrchov na území Slovenska a Česka, je takmer nemožné. Problematickosť stratifikovaného prístupu pri potrebe dosiahnutia spoľahlivého odhadu chyby vynechania priznávajú aj Sannier et al. (2016, p. 64 – 66), ktorí navrhujú hodnotiť tematickú presnosť IMD2006 až IMD2012 pomocou stratifikovaného systematického výberu s využitím databázy LUCAS (Land Use and Coverage Area frame Survey) s celkovým počtom 20 164 štvorcových vzoriek s rozmermi 100 × 100 m (každá s pravidelnou sieťou 5 × 5 interpretovaných bodov s rozstupmi 20 m) v 39 krajinách EEA (teda asi niekoľko stoviek štvorcových vzoriek na Slovensku a v Česku spolu), a ktorých návrh bol odvtedy aplikovaný aj pri hodnotení IMD2015 a IMD2018 s uvedením zmienky o tomto probléme (Smith 2020, p. 18). Riešením je použitie náhodného výberu s výrazne väčším počtom vzoriek. Príkladom takéhoto prístupu je práca Krówczyńskiej et al. (2016).

Prvé tri práce spomenuté v predchádzajúcom odseku hodnotiace tematickú presnosť IMD však neanalyzujú priamo pôvodnú vrstvu s pomerovou (metrickou) premennou (s celočíselnými hodnotami 0 až 100 %), ale transformovanú vrstvu s dichotomickou premennou (s hodnotami „zastavaná“ a „nezastavaná“) odvodenú

z pôvodnej vrstvy pomocou aplikácie jedného alebo viacerých prahov. Gallego et al. (2016, p. 3) však navrhujú vychádzať pri hodnotení priamo z celočíselných hodnôt 0 až 100 % v jednotlivých 100-metrových pixeloch reprezentujúcich percentuálny podiel nepriepustných povrchov na celkovej ploche daného pixela, pričom využívajú vyššie spomenutý stratifikovaný systematický výber na báze LUCAS.

ZÁUJMOVÉ ÚZEMIE A DÁTA

Keďže cieľom príspevku je analyzovať vývoj tematickej presnosti doteraz zverejnených verzií vrstvy hustoty nepriepustnosti IMD pre roky 2006 a 2009 na Slovensku a v Česku, je potrebné tieto verzie špecifikovať. Podarilo sa nám analyzovať 4 z 5 postupne zverejňovaných aktualizáčnych verzií IMD2006 a 2 z 3 postupne zverejňovaných aktualizáčnych verzií IMD2009:

- IMD2006_v1 – dostupná od 12. 6. 2009, tzv. „precursor“,
- IMD2006_v2 – dostupná od 26. 1. 2010, tzv. „enhancement“,
- IMD2006_v3, IMD2009_v1 – dostupné od 22. 1. 2013, publikované pri príležitosti vytvorenia IMD2009,
- IMD2006_v4, IMD2009_v2 – dostupné približne v roku 2014, publikované pri príležitosti vytvorenia IMD2012 (v čase dostupnosti sa nám ich nepodarilo získať, pričom EEA sa rozhodla neposkytovať verejnosti staršie verzie týchto produktov po zverejnení ich najnovších verzií),
- IMD2006_v5, IMD2009_v3 – dostupné od 15. 1. 2019, publikované pri príležitosti vytvorenia IMD2015.

Na hodnotenie tematickej presnosti týchto verzií IMD bol použitý najčastejšie využívaný postup validácie produktov vo forme rastra s kvantitatívnymi hodnotami, akým je aj IMD, ktorý Hurbánek a Rosina (2013) stručne popísali v nasledovných krokoch:

- vytvorí sa jednoduchý náhodný výber dostatočne veľkej vzorky pixelov spomedzi všetkých pixelov mapy,
- zistia sa referenčné hodnoty skúmanej premennej v pixeloch vybranej vzorky,
- porovnajú sa mapové a referenčné hodnoty danej vzorky,
- kvantifikuje sa „(ne)zhoda“ medzi týmito dvojicami hodnôt danej vzorky.

V zmysle prvých dvoch krokov bol preto na účely tejto analýzy vytvorený náhodný výber 20 000 vzoriek pixelov s rozlíšením 100 m tak na území Slovenska, ako aj na území Česka. Vzorky boli náhodne vyberané len z pixelov s hodnotami 0 až 100 %, teda nie aj z pixelov s hodnotami 254 (neklasifikovateľné, napr. kvôli pokrytiu oblakov) alebo 255 (bez dát, napr. hraničný pixel). V každom pixeli vo vzorke bol následne uplatnený systematický výber 100 bodov v pravidelnej matici 10 × 10 bodov s rozostúpením 10 m.

Zvolený počet 20 000 vzoriek je na jednej strane kompromisom medzi snahou o maximalizáciu presnosti odhadu percentuálneho podielu nepriepustných povrchov z celkovej rozlohy danej krajiny v danom roku (nižšie označovaného P), resp. minimalizáciu šírky príslušného intervalu spoľahlivosti (ktorá sa znižuje s rastúcim počtom vzoriek), a na druhej strane snahou o minimalizáciu náročnosti tvorby referenčnej databázy (ktorá sa znižuje s klesajúcim počtom vzoriek).

Vizuálna interpretácia snímok prebiehala v programe ArcMap 10.2.2 v jednotlivých pixeloch „na slepo“ – hodnotiaci osoba nepoznala hodnotu nepriepustnosti pixela vo validovanej databáze, aby nebol úsudok ovplyvnený. V prípade Česka bola síce takáto náhodne vytvorená vzorka pre rok 2009 interpretovaná už skôr (Pohlová 2013), avšak aby bola zachovaná konzistentnosť interpretácie s rokom 2006 v Česku a s obidvoma rokmi na Slovensku, v rámci tejto štúdie bola reinterpretovaná.

Každému bodu bola priradená jedna z nasledujúcich troch hodnôt podľa toho, na akom povrchu sa tento bezrozmerný bod na ortofotosnímke pre daný rok nachádzal:

0 priepustný povrch,

1 nepriepustný povrch s alochtónnym materiálom (premiestneným na dané miesto aktivitou človeka) – napr. budovy, cesty a parkoviská,

2 nepriepustný povrch s autochtónnym materiálom („znepriepustnená“ vďaka intenzívnej aktivite človeka, indikovaná absenciou vegetácie) – napr. poľné cesty.

Povrchy boli pritom interpretované tak, ako boli nasnímané. Napríklad, ak sa v lokalite zosnímanej pod iným ako pravým uhlom vyskytol bod umiestnený na vertikálnej ploche, napr. na múre výškovej budovy, tento bod bol interpretovaný ako bod nepriepustný s hodnotou 1. Tento postup vychádzal z predpokladu, že snímky, na báze ktorých bola IMD generovaná, nemali charakter pravých, resp. skutočných ortofotosnímkov (tzv. true ortho), a preto by ho nemali mať ani referenčné snímky. Naopak, ak bol bod umiestnený na korune stromu, pod ktorým sa – súdiac z kontextu – zjavne nachádzal nepriepustný povrch, bola mu priradená hodnota 0, bol teda považovaný za priepustný. Interpretujúca osoba v takomto prípade nevyužila dostupnú kontextuálnu informáciu rovnako, ako ju neskúmal a nevyužíval ani algoritmus generovania IMD. Pri vyhodnocovaní slabo viditeľných povrchov nachádzajúcich sa v tieni však kontextuálne informácie boli využité. Medzi ďalšie zaujímavé príklady patria ťažko identifikovateľné povrchy malých rozmerov, často lokalizované v blízkosti budov, na dvoroch, ako napríklad auto, vrak auta, klietka pre domáce zvieratá, kopa štrku, dreva alebo iného materiálu, keď bolo zjavné, že sa pod nimi nachádza priepustný povrch. Preto bolo potrebné skúmať tieto objekty z hľadiska dĺžky ich prítomnosti na danom mieste, a to porovnaním stavu danej lokality na iných snímkach príslušného alebo časovo relatívne blízkeho obdobia. Ak bol objekt považovaný za krátkodobu lokalizovaný na danom priepustnom povrchu (pretože na iných snímkach sa nevyskytoval alebo sa vyskytoval v inej polohe), bola mu priradená hodnota 0. Naopak, ak bol objekt považovaný za dlhodobu lokalizovaný na priepustnom povrchu, dostal hodnotu 1. Špecifické prípady popísané v tomto odseku mali však len minimálny vplyv na celkovú zistenú veľkosť a charakter tematickej presnosti.

Pre popísanie nejednoznačnosti definície pojmu nepriepustnosť boli vytvorené dve verzie referenčných databáz (IR):

IR_A – body s hodnotou 2 boli považované za priepustné povrchy (teda rovnako ako 0),

IR_B – body s hodnotou 2 boli považované za nepriepustné povrchy (teda rovnako ako 1).

Spočítaním bodov nachádzajúcich sa na nepriepustnom povrchu bola vyjadrená referenčná hodnota podielu nepriepustných povrchov každého pixelu (od 0 do 100 %), keďže každý pixel obsahoval presne 100 bodov.



Obr. 1. Ukážka jedného pixela 100×100 m vo vzorke a rozmiestnenia interpretovaných bodov v ňom s rozstupmi 10 m v prostredí ArcMap 10.2.2. Príslušné hodnoty tejto vzorky (v %) v IMD_2006_v1 až v5 sú 77, 77, 68, 27, v IR_2006_A 53 a v IR_2006_B 57.

Zdroj dát: Eurosense a Geodis Slovakia (2007).

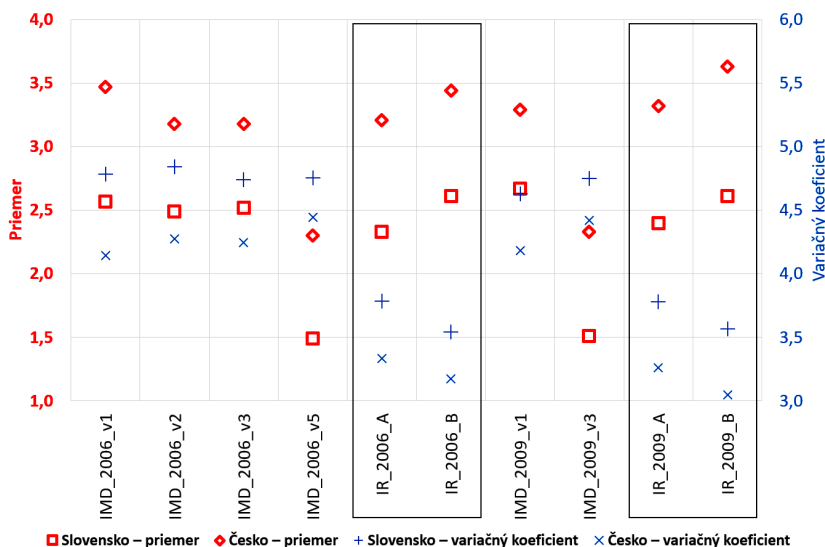
Na Slovensku boli použité predovšetkým letecké ortofotosnímky nasnímané v rokoch 2005 až 2007 a 2008 až 2010 (Eurosense a Geodis Slovakia 2007 a 2010). Zatiaľ čo tie staršie boli dostupné online, napr. aj cez Google Maps alebo Google Earth, tie novšie bolo potrebné zakúpiť. Pre ich vysokú cenu bolo možné zakúpiť ortofotosnímky len pre 5 000 vzoriek, v ktorých sa ich nulový podiel nepriepustných povrchov medzi rokmi 2006 a 2009 s najväčšou pravdepodobnosťou nezmenil (zväčša lokality značne vzdialené od existujúcich nepriepustných povrchov). Na Slovensku tak bolo pre referenčný rok 2006 interpretovaných 20 000 pixelov a pre referenčný rok 2009 len 5 000 zakúpených vybraných pixelov (pričom zvyšným 15 000 pixelom bola priradená nezmenená referenčná hodnota intenzity nepriepustnosti 0 %). V Česku boli použité predovšetkým letecké ortofotosnímky nasnímané v rokoch 2005 až 2007 a 2008 až 2010 (ČÚZK 2022a) dostupné cez príslušnú WMS službu (ČÚZK 2022b). Interpretovaných bolo 20 000 pixelov pre každý referenčný rok.

Vzhľadom na možnú časovú nezrovnalosť medzi momentom vytvorenia satelitných snímok použitých pri produkcii IMD a momentom vytvorenia uvedených leteckých ortofoto-snímok boli v niektorých špecifických prípadoch využité aj iné online verejne dostupné produkty DPZ.

METODIKA A VÝSLEDKY

Každá verzia databázy vrstvy hustoty nepriepustnosti IMD bola porovnaná s príslušnou referenčnou databázou IR pre obidva roky a pre obidve krajiny na úrovni jednotlivých 100-metrových pixelov. Po úvodnom porovnaní základných deskriptívno-štatistických charakteristík IMD a IR, z ktorých je na prvý pohľad zrejmé, že najnovšie verzie IMD sa v porovnaní s prechádzajúcimi verziami svojím

priemerom¹ výraznejšie vzdialili od priemeru príslušnej referenčnej databázy (obr. 2), boli na hodnotenie vývoja tematickej presnosti použité tieto hodnotiace nástroje: histogramy hodnôt IMD, IR a ich rozdielov IMD – IR (chýb), korelačná analýza vzťahov medzi IR a IMD, celková absolútna chyba relativizovaná (TAER) a jej štruktúra a regresná analýza vzťahov medzi IR a rozdielmi IMD – IR (chybami).



Obr. 2. Priemer a variačný koeficient jednotlivých verzií databázy vrstvy hustoty nepriepustnosti IMD a referenčnej databázy IR

Zdroj dát: autori.

Histogramy

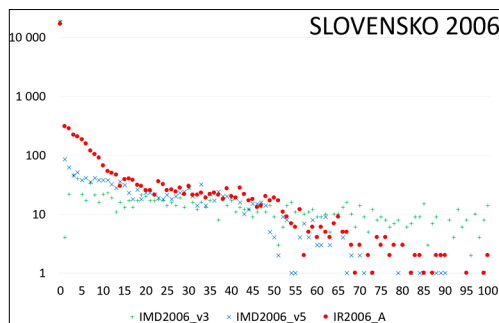
Hodnotenie pomocou histogramov² je vhodné na porovnanie početnosti jednotlivých hodnôt hustoty nepriepustnosti v pôvodnej databáze IMD a v referenčnej databáze IR (obr. 3 až 6). V ideálnom prípade absolútnej tematickej presnosti by boli histogramy príslušných dvojíc databáz zhodné.

Početnosť takmer všetkých nenulových hodnôt nepriepustnosti (skoro všetkých menších hodnôt v oboch krajinách a väčšiny väčších hodnôt na Slovensku) sa zdá byť podhodnotená vo všetkých najnovších verziách IMD (krivka spájajúca modré krížiky je takmer v celej dĺžke pod krivkou spájajúcou červené bodky v obr. 3 až 6), zatiaľ čo frekvencia nulových hodnôt nepriepustnosti je nadhodnotená. Avšak v predchádzajúcich dostupných verziách pre obidva roky (IMD2006_v3 a

¹ Aritmetický priemer všetkých 20 000 hodnôt pixelov vo vzorke vyjadruje priemerný percentuálny podiel nepriepustných povrchov z celkovej rozlohy záujmového územia podľa príslušnej databázy IMD a IR. Všetky priemery spomínané v článku sú aritmetické. Priemer referenčnej databázy, nižšie označovaný P a príslušné 95 % intervaly spoľahlivosti pre IR_A sú pre Slovensko $P_{2006} = 2,33\%$ (2,21 a 2,45) a $P_{2009} = 2,40\%$ (2,27 a 2,53) a pre Česko $P_{2006} = 3,21\%$ (3,06 a 3,36) a $P_{2009} = 3,32\%$ (3,17 a 3,47).

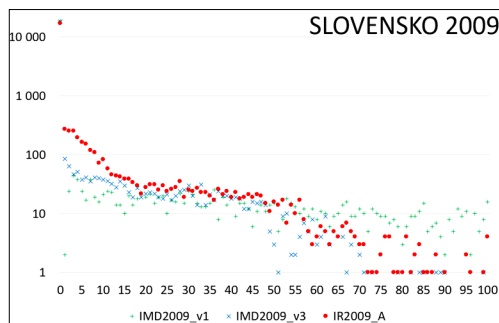
² Je zrejme, že histogramy sú definované ako stĺpcové grafy. Histogramy v tomto príspevku sú zámerné transformované do bodových grafov (pričom bod vystihuje najvyššiu hodnotu daného stĺpca), čo umožňuje detailné a názorné porovnanie rozdelenia dvoch a viacerých súborov aj pri použití 101 intervalov so šírkou jedného percentuálneho bodu, teda 101 tried na osi x: 0, 1, 2,... 100 %.

IMD2009_v1) sa početnosť väčších hodnôt nepriepustnosti zdá byť nadhodnotená, pričom toto nadhodnotenie je na Slovensku o niečo výraznejšie.



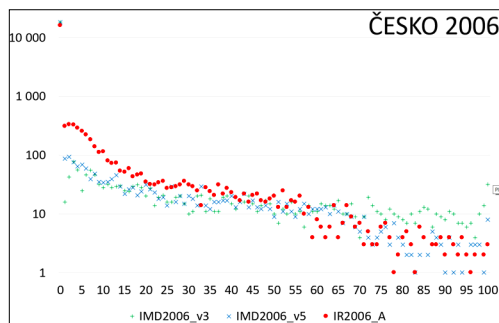
Obr. 3. Histogram rozdelenia hodnôt hustoty nepriepustnosti v databázach IMD2006_v3 a IMD2006_v5 a v referenčnej databáze IR2006_A pre Slovensko

Zdroj dát: autori.



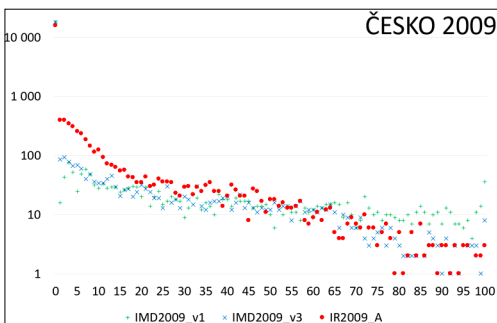
Obr. 4. Histogram rozdelenia hodnôt hustoty nepriepustnosti v databázach IMD2009_v1 a IMD2009_v3 a v referenčnej databáze IR2009_A pre Slovensko

Zdroj dát: autori.



Obr. 5. Histogram rozdelenia hodnôt hustoty nepriepustnosti v databázach IMD2006_v3 a IMD2006_v5 a v referenčnej databáze IR2006_A pre Česko

Zdroj dát: autori.



Obr. 6. Histogram rozdelenia hodnôt hustoty nepriepustnosti v databázach IMD2009_v1 a IMD2009_v3 a v referenčnej databáze IR2009_A pre Česko

Zdroj dát: autori.

Okrem histogramov reprezentujúcich rozdelenie početností samotných hodnôt 0, 1, 2, ..., 100 % jednotlivých verzií databáz IR a IMD boli pre lepšie pochopenie vývoja tematickej presnosti zostrojené tiež histogramy početností hodnôt chýb, t. j. rozdielov dvojíc hodnôt IMD a IR v jednotlivých vzorkách tak, ako to navrhujú napr. Gallego et al. (2016, p. 3). Vypočítané chyby nadhodnotenia a podhodnotenia v jednotlivých vzorkách databáz sú na obr. 7 rozdelené na dve skupiny (podľa tab. 2): chyby závažnejšieho podhodnotenia (MaU) a závažnejšieho nadhodnotenia (MaO) sú vyznačené červenou farbou a chyby menej závažného podhodnotenia (MiU) a menej závažného nadhodnotenia (MiO) – spolu so vzorkami s nulovým rozdielom medzi IMD a IR – sú vyznačené modrou farbou.

Klasifikácia chýb uvedená v tab. 2, ako aj termíny „závažnejšie“ a „menej závažné“ použité v názvoch príslušných typov chýb, boli zvolené s ohľadom na hlav-

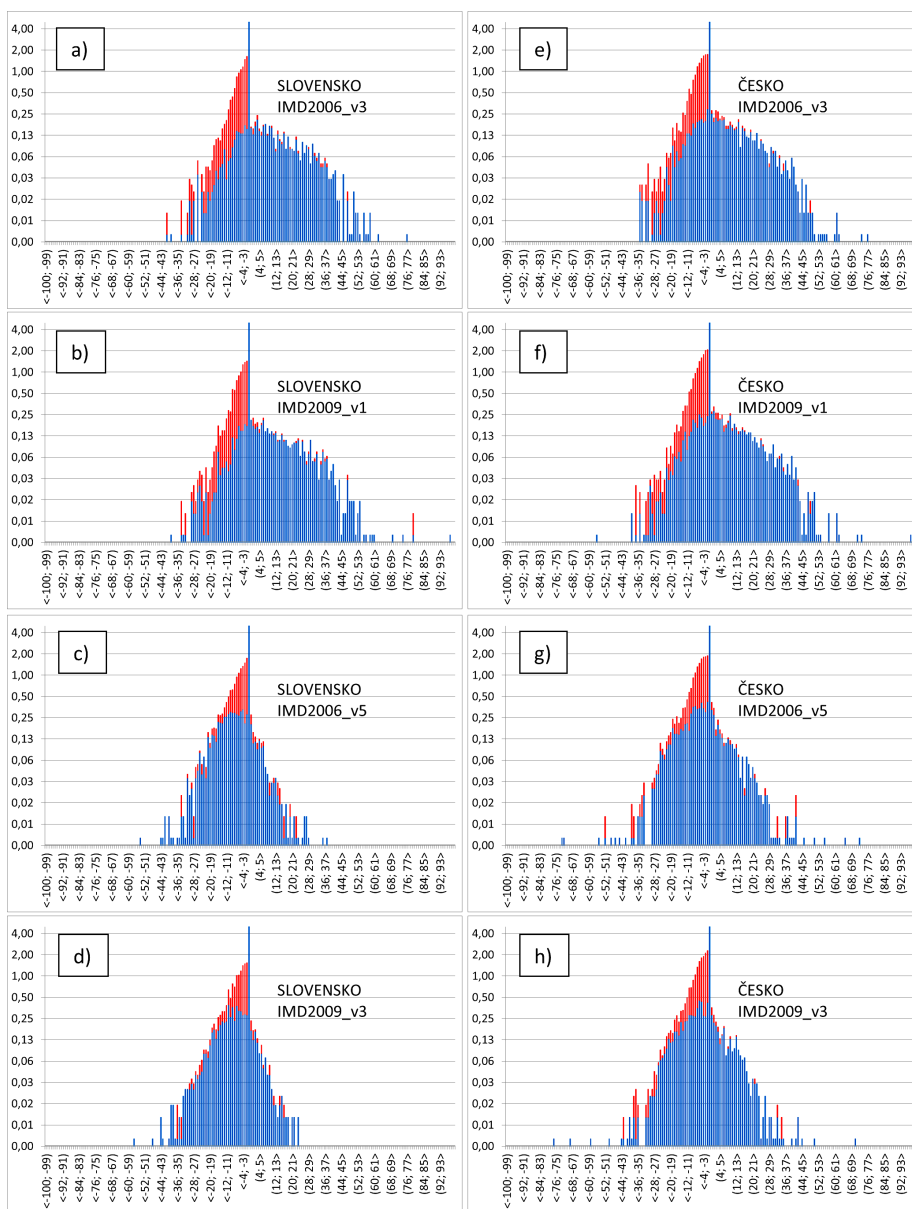
ný cieľ práce. Pri inom ciele by bolo legitímne použiť inú klasifikáciu s inými názvami typov chýb. Pre úplnosť je vhodné dodať, že chyby typu MaO a MaU sú v tejto práci chápané de facto rovnako ako sú chyby priradenia a vynechania chápané v literatúre venovanej téme validácie kvalitatívnych (kategórických) tematických dátových vrstiev, resp. kvantitatívnych vrstiev transformovaných pomocou zvoleného prahu na kvalitatívne. Zatiaľ čo napríklad v tab. 1 alebo v práci Mauchu a Büttnera (2008) bola zvolená prahová hodnota podielu nepriepustných povrchov 80 %, v práci Mauchu et al. (2010) to bolo 30 % a v klasifikácii v tab. 2 to je – pri celočíselných hodnotách – 1 % (nie však pre celkové a jediné hodnotenie IMD, ale len pre časť hodnotenia venovanú histogramom chýb a štruktúre celkovej absolútnej chyby relativizovanej TAER).

Tab. 2. Typy pixelov podľa rozdielu medzi validovanou a príslušnou referenčnou hodnotou

| IMD hodnoty | IR hodnoty | Rozdiel | Klasifikácia pixelov |
|-------------|------------|----------------|--|
| IMD_i | IR_i | $IMD_i - IR_i$ | Typy pixelov |
| 0 | 0 | 0 | AP – správne klasifikované ako priepustné agreement pervious |
| > 0 | > 0 | 0 | AI – správne klasifikované ako nepriepustné agreement impervious |
| > 0 | > 0 | > 0 | MiO – menej závažné nadhodnotenie minor overestimation |
| > 0 | > 0 | < 0 | MiU – menej závažné podhodnotenie minor underestimation |
| > 0 | 0 | > 0 | MaO – závažnejšie nadhodnotenie major overestimation |
| 0 | > 0 | < 0 | MaU – závažnejšie podhodnotenie major underestimation |

Zdroj: autori.

Zatiaľ čo záporné hodnoty -100 až -1 na osi x reprezentujú vzorky s chybami podhodnotenia ($IMD_i - IR_i < 0$), kladné hodnoty 1 až 100 na osi x predstavujú vzorky s chybami nadhodnotenia ($IMD_i - IR_i > 0$). Týchto dvesto možných stĺpcov v každom histograme na obr. 7 dopĺňa 201. stĺpec pre nulovú „chybu“ ($IMD_i - IR_i = 0$), ktorá je najpočetnejšia. Prvý stĺpec histogramov (a, b, c, d) zachytáva početnosti chýb na Slovensku a druhý stĺpec histogramov početnosti chýb (e, f, g, h) v Česku. Štyri horné histogramy (a, b, e, f) znázorňujú predposledné dostupné verzie vrstvy hustoty nepriepustnosti pre roky 2006 a 2009 (IMD_{2006_v3} a IMD_{2009_v1}) a štyri dolné histogramy (c, d, g, h) ich najnovšie verzie (IMD_{2006_v5} a IMD_{2009_v3}).



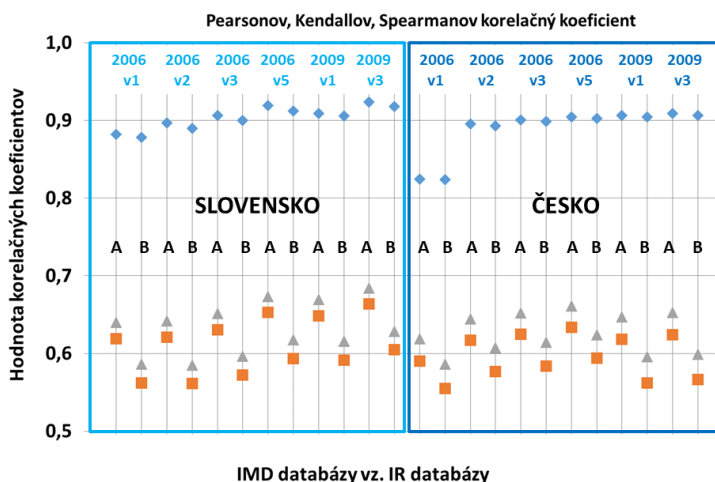
Obr. 7. Histogramy relatívnych početností rôznych hodnôt chýb nadhodnotenia a podhodnotenia získané porovnaním IMD2006_v3, IMD2006_v5, IMD2009_v1, IMD2006_v3 s príslušnými databázami IR2006_A a IR2009_A pre Slovensko a Česko. Značky na osi y (v percentách z celkového počtu 20 000 pixelov) s logaritmickou mierkou pri 11 líniách horizontálnej mriežky zhora nadol: 4; 2; 1; 0,5; 0,25; 0,125; 0,0625; 0,03125; 0,015625; 0,0078125; 0,00390625. Značky na osi x (v percentuálnych bodoch) pri jednotlivých stĺpcoch zľava doprava: -100, -99, ..., -2, -1, 0, 1, 2, ..., 99, 100.

Zdroj dát: autori.

Vo všetkých histogramoch na obr. 7 je viac podhodnotených pixelov s menšími chybami podhodnotenia ako nadhodnotených pixelov s menšími chybami nadhodnotenia. V starších verziách IMD (a, b, e, f) je viac nadhodnotených pixelov s väčšími chybami nadhodnotenia ako podhodnotených pixelov s väčšími chybami podhodnotenia. Platí teda, že sa v týchto starších verziách vyskytuje relatívne veľa malých chýb podhodnotenia a relatívne málo veľkých chýb nadhodnotenia. Pritom, z pohľadu potenciálneho využitia IMD pre dezagregáciu sociálnoekonomických premenných (v zmysle tab. 2), pri chybách podhodnotenia prevažujú závažnejšie chyby a pri chybách nadhodnotenia menej závažné chyby v obidvoch rokoch a v obidvoch krajinách. V najnovších verziách IMD (c, d, g, h) je v porovnaní s tými predchádzajúcimi zrejmy pokles počtu väčších chýb nadhodnotenia a tiež pokles počtu závažnejších (najmä stredne veľkých) chýb podhodnotenia, opäť v obidvoch rokoch a v obidvoch krajinách.

Korelačná analýza

Na hodnotenie tematickej presnosti vrstvy hustoty nepriepustnosti je možné využiť aj Pearsonov, Kendallov a Spearmanov korelačný koeficient. Korelačná analýza kvantifikuje tesnosť a charakter vzťahu medzi referenčnými a validovanými dátami. Čím viac sa hodnota koeficientu približuje k hodnote 1,0, tým je tematická presnosť IMD väčšia. Vzhľadom na skutočnosť, že vstupné dáta nemajú normálne rozdelenie (o. i. obsahujú veľmi veľký počet pixelov s nulovou hodnotou, ktoré nie sú z korelačnej analýzy vylúčené, podobne ako napr. v práci Smitha (2020, p. 45), získané hodnoty Pearsonovho korelačného koeficientu majú, na rozdiel od zvyšných dvoch koeficientov, len obmedzenú výpovednú hodnotu. Sú však uvádzané pre porovnanie, keďže štandardne sa pri hodnotení IMD využívajú, a to aj bez zmienky príslušných hodnôt intervalov spoľahlivosti (Smith 2020, p. 50).



Obr. 8. Hodnoty korelačných koeficientov pre vzťahy medzi referenčnými databázami IR a príslušnými validovanými databázami IMD pre roky 2006 a 2009 na Slovensku a v Česku

Zdroj dát: autori.

Podľa obr. 7 sa zdá byť referenčná databáza IR_A presnejšia referenčná databáza IR_B, keďže hodnoty korelačných koeficientov verzie B sú vo všetkých prípadoch nižšie v porovnaní s príslušnými hodnotami korelačných koeficientov verzie A. Tematická presnosť postupne narastá od starších verzií k novším. Vzhľadom na referenčnú databázu IR_A je IMD na Slovensku tematicky presnejšia ako v Česku pre obidva roky v takmer všetkých verziách a podľa bežných všetkých koeficientov (výnimkou sú verzie IMD2006_v2 a IMD2006_v3, a to len podľa Spearmanovho koeficientu). Vzhľadom na referenčnú databázu IR_B je IMD na Slovensku tematicky presnejšia ako v Česku pre rok 2009 v obidvoch verziách a podľa všetkých koeficientov, no vo verzii IMD2006_v2 je to podľa všetkých koeficientov naopak a v ostatných verziách pre rok 2006 výsledok porovnania závisí od použitého koeficientu.

Celková absolútna chyba relativizovaná (TAER) a jej štruktúra

Pri validácii intenzitných máp sa obyčajne vypočítavajú hodnoty celkovej absolútnej chyby (TAE). Nevýhodou ukazovateľa TAE je skutočnosť, že pomocou nej nemožno porovnávať tematickú presnosť máp dvoch rôznych území, ktoré majú rôzny reálny podiel nepriepustných povrchov (Hurbánek a Rosina 2013). Ako riešenie sa ponúka možnosť relativizovať výslednú TAE (príp. priemernú absolútnu chybu TAE/n) hodnotou najlepšieho dostupného odhadu rozlohy (príp. podielu) nepriepustných povrchov v záujmovom území, za ktorý možno považovať pri dostatočne veľkej náhodnej vzorke sumu (príp. priemer) referenčných hodnôt všetkých pixelov vo vzorke (P):

$$P = \frac{\sum_i^n IR_i}{n}$$

Výsledkom takejto relativizácie je celková absolútna chyba relativizovaná (TAER):

$$TAE = \sum_i^n |IMD_i - IR_i| \quad TAER = \frac{\sum_i^n |IMD_i - IR_i|}{\sum_i^n IR_i} \times 100(\%) = \frac{TAE}{P} \times 100(\%),$$

kde IMD_i je mapová hodnota podielu nepriepustných povrchov v i -tom pixeli (%), IR_i je referenčná hodnota podielu nepriepustných povrchov v i -tom pixeli (%).

Okrem hodnotenia celkovej veľkosti TAER je veľmi dôležité aj hodnotenie jej štruktúry. Vzniknuté chyby nadhodnotenia a podhodnotenia v pixeloch boli rozdelené na dva typy – menej závažné a závažnejšie, z pohľadu ich potenciálneho vplyvu na výsledok pri využívaní IMD ako pomocnej vrstvy pre dezagregáciu sociálno-ekonomických premenných (tab. 2).

Na tento účel je totiž potrebné poznať nielen celkovú chybu, ale tiež to, ako často sa tieto dva typy zhody a štyri typy nezahody (resp. chýb) vyskytujú v IMD. Pre každý z týchto šiestich typov (ne)zhody je preto potrebné zistiť frekvenciu jeho výskytu (n_{AP} , n_{AI} , n_{MiO} , n_{MiU} , n_{MaO} , n_{MaU}) vzhľadom na celkovú veľkosť vzorky (n), pričom platí:

$$n_{AP} + n_{AI} + n_{MiO} + n_{MiU} + n_{MaO} + n_{MaU} = n$$

$$\frac{n_{AP}}{n} + \frac{n_{AI}}{n} + \frac{n_{MiO}}{n} + \frac{n_{MiU}}{n} + \frac{n_{MaO}}{n} + \frac{n_{MaU}}{n} = 100\%$$

Rovnako dôležité je poznať tiež veľkosť príspevku každého zo štyroch typov nezahody k celkovej hodnote chyby TAER, a teda do akej miery sa každý z nich podieľa na celkovej tematickej nepresnosti IMD. Pre hodnoty jednotlivých štyroch typov nezahody (a dvoch typov zhody) na celkovej hodnote chyby TAER pritom platí:

$$TAER = TAER_{MiO} + TAER_{MaO} + TAER_{MiU} + TAER_{MaU} \quad (TAER_{AP} = 0; TAER_{AI} = 0)$$

$$\text{pre MiO: } TAER_{MiO} = \frac{\sum_i^{n_{MiO}} |IMD_i - IR_i|}{\sum_i^n IR_i} \cdot 100(\%)$$

$$\text{pre MaO } TAER_{MaO} = \frac{\sum_i^{n_{MaO}} |IMD_i - IR_i|}{\sum_i^n IR_i} \cdot 100(\%)$$

$$\text{pre MiU } TAER_{MiU} = \frac{\sum_i^{n_{MiU}} |IMD_i - IR_i|}{\sum_i^n IR_i} \cdot 100(\%)$$

$$\text{pre MaU } TAER_{MaU} = \frac{\sum_i^{n_{MaU}} |IMD_i - IR_i|}{\sum_i^n IR_i} \cdot 100(\%)$$

Obdobne definovaná celková absolútna chyba relativizovaná pre chyby podhodnotenia ($TAER_U$) je zhodná s chybou vynechania tak, ako ju chápu Gallego et al. (2016, p. 3), Sannier et al. (2016, p. 64) alebo Smith (2020, p. 24):

$$\text{pre U (MiU+MaU): } TAER_U = \frac{\sum_i^{n_U} |IMD_i - IR_i|}{\sum_i^n IR_i} \cdot 100(\%)$$

Ak by v menovateli výpočtu celkovej absolútnej chyby relativizovanej pre chyby nadhodnotenia ($TAER_O$) bola suma IMD_i namiesto sumy IR_i , bol by tento výpočet zhodný s výpočtom chyby priradenia uvádzaným tými istými autormi:

$$\text{pre O (MiO+MaO): } TAER_O = \frac{\sum_i^{n_O} |IMD_i - IR_i|}{\sum_i^n IR_i} \cdot 100(\%)$$

V tab. 3 a obr. 9 sú uvedené hodnoty TAER pre porovnanie s referenčnou databázou IR_A, keďže sa už vyššie sa ukázalo, že táto vykazuje väčšiu zhodu s príslušnými verziami IMD. Pre výpočet hodnôt v km² bola použitá celková rozloha Slovenska 49 036 km² a Česka 78 866 km². Z hodnôt TAER vyplýva, že rozloha nesprávne klasifikovaných plôch v pomere k rozlohe nepriepustných povrchov v príslušnom záujmovom území tvorí od 44,36 % (na Slovensku v najnovšej verzii IMD2009) až do 70,73 % (v Česku v najstaršej verzii IMD2006). V najstaršej verzii IMD2006 a v najnovších verziách IMD2006 a IMD2009 bola hodnota TAER väčšia v Česku v porovnaní so Slovenskom, v ostatných verziách to bolo naopak.

V tabuľke 3 stĺpec 2 (TAE/n) vyjadruje podiel nesprávne klasifikovaných plôch z celkovej rozlohy danej krajiny v danom roku, teda zo súčtu rozlohy tak priepustných, ako aj nepriepustných povrchov v danej krajine v danom roku. Keďže rozlo-

ha nepriepustných povrchov sa v jednotlivých krajinách pre jednotlivé roky líši, na porovnanie tematickej presnosti naprieč krajinami a rokmi je vhodnejšie použiť stĺpec 4 (TAER) vyjadrujúci pomer rozlohy nesprávne klasifikovaných plôch k rozlohe nepriepustných povrchov v danej krajine v danom roku.

Tab. 3. Celková absolútna chyba relativizovaná (TAER) pre referenčnú databázu IR_A

| Stĺpec 1 | Stĺpec 2 (%) | Stĺpec 3 (km ²) | Stĺpec 4 (%) | Stĺpec 5 (%) | Stĺpec 6 (km ²) |
|------------------------|--------------|-----------------------------|--------------|--------------|-----------------------------|
| SK_IMD2006_v1_IR2006_A | 1,54 | 753,59 | 65,95 | 2,33 | 1 142,71 |
| SK_IMD2006_v2_IR2006_A | 1,47 | 720,71 | 63,07 | 2,33 | 1 142,71 |
| SK_IMD2006_v3_IR2006_A | 1,41 | 691,43 | 60,51 | 2,33 | 1 142,71 |
| SK_IMD2006_v5_IR2006_A | 1,06 | 517,57 | 45,29 | 2,33 | 1 142,71 |
| SK_IMD2009_v1_IR2009_A | 1,44 | 705,90 | 59,89 | 2,40 | 1 178,70 |
| SK_IMD2009_v3_IR2009_A | 1,07 | 522,92 | 44,36 | 2,40 | 1 178,70 |
| CR_IMD2006_v1_IR2006_A | 2,27 | 1 791,24 | 70,73 | 3,21 | 2 532,39 |
| CR_IMD2006_v2_IR2006_A | 1,86 | 1 469,43 | 58,03 | 3,21 | 2 532,39 |
| CR_IMD2006_v3_IR2006_A | 1,82 | 1 434,73 | 56,66 | 3,21 | 2 532,39 |
| CR_IMD2006_v5_IR2006_A | 1,48 | 1 170,45 | 46,22 | 3,21 | 2 532,39 |
| CR_IMD2009_v1_IR2009_A | 1,87 | 1 476,49 | 56,40 | 3,32 | 2 618,04 |
| CR_IMD2009_v3_IR2009_A | 1,52 | 1 200,18 | 45,84 | 3,32 | 2 618,04 |

Stĺpec 1 – Názov konkrétnej dvojice porovnávaných verzií validovanej databázy IMD a referenčnej databázy IR;
Stĺpec 2 – Priemerná absolútna chyba (resp. priemer absolútnych hodnôt chýb), t. j. odhad podielu nesprávne klasifikovaných plôch z rozlohy danej krajiny v danom roku $TAE/n = (\sum |IMDi - IRi|) / n$ (%);

Stĺpec 3 – Odhad rozlohy nesprávne klasifikovaných plôch v danej krajine v danom roku, t. j. predchádzajúci stĺpec vynásobený rozlohou krajiny (km²);

Stĺpec 4 – Priemerná absolútna chyba (stĺpec 2) v pomere k priemernej referenčnej hodnote (stĺpec 5), resp. odhad rozlohy nesprávne klasifikovaných plôch (stĺpec 3) v pomere k odhadu rozlohy nepriepustných povrchov (stĺpec 6), t. j. celková absolútna chyba relativizovaná v danej krajine v danom roku $TAER = (TAE/n) / P \cdot 100$ (%);

Stĺpec 5 – Priemerná referenčná hodnota, t. j. odhad podielu nepriepustných povrchov z rozlohy danej krajiny v danom roku $P = (\sum IRi) / n$ (%);

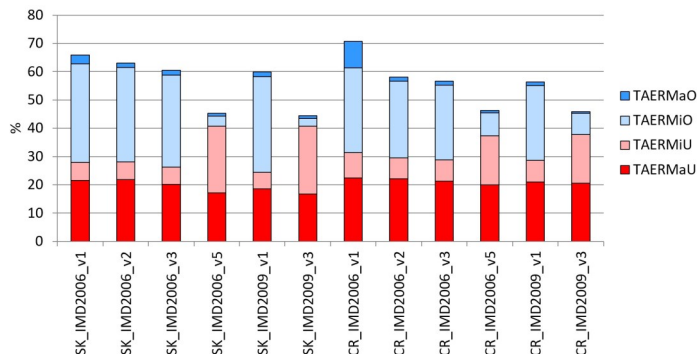
Stĺpec 6 – Odhad rozlohy nepriepustných povrchov v danej krajine, t. j. predchádzajúci stĺpec vynásobený rozlohou krajiny (km²).

Poznámka: Rozloha krajiny použitá vo výpočte je 49 036 km² pre Slovensko a 78 866 km² pre Česko.

Zdroj dát: autori.

Z obr. 9 je zrejmé, že vypočítaná chyba TAER klesá v oboch krajinách pre obidva skúmané roky od starších verzií po novšie, a teda že tematická presnosť jednotlivých verzií v čase postupne narastá. K najvýraznejšiemu zlepšeniu pritom došlo medzi dvoma najnovšími dostupnými verziami, t. j. medzi IMD2006_v3 a IMD2006_v5 pre vrstvu pre rok 2006 (zo 60,51 % na 45,29 % na Slovensku a z 56,66 % na 46,22 % v Česku) a medzi IMD2009_v1 a IMD2009_v3 pre vrstvu pre rok 2009 (z 59,89 % na 44,36 % na Slovensku a z 56,40 % na 45,84 % v Česku). Medzi uvedenými dvojicami verzií IMD však zároveň došlo k výraznej zme-

ne štruktúry TAER, v porovnaní so zmenami štruktúry pri predchádzajúcich vylepšeniach IMD. Hoci v najnovších verziách je podiel menej závažných chýb nadhodnotenia výrazne menší, podiel menej závažných chýb podhodnotenia je výrazne väčší ako v predchádzajúcich verziách.



Obr. 9. Celková absolútna chyba relativizovaná (TAER) v percentách a jej štruktúra vypočítaná pre referenčnú databázu IR_A

Zdroj dát: autori.

Pri niektorých aplikáciách, ako napr. pri dezagregácii sociálno-ekonomických premenných, sú za najkritickejšie považované závažnejšie chyby ($TAER_{Ma} = TAER_{MaO} + TAER_{MaU}$). Žiaľ čo v procese dezagregácie závažnejšie chyby nadhodnotenia ($TAER_{MaO}$) vytvárajú fiktívne dazymetrické zóny s relatívne veľkou predpokladanou hustotou zaľudnenia, naopak závažnejšie chyby podhodnotenia ($TAER_{MaU}$) neodhaľujú skutočné dazymetrické zóny s relatívne veľkou predpokladanou hustotou zaľudnenia. Je pozitívne, že veľkosť týchto závažnejších chýb v smere od starších k novším verziám klesá, avšak pre pomalé tempo tohto poklesu (vzhľadom na rýchlejší pokles celkovej chyby TAER) ich relatívny podiel na celkovej rozlohe nesprávne klasifikovaných plôch v čase narastá, napr. podiel $TAER_{MaU}$ v najnovších verziách na takmer 38 % na Slovensku a 43 až 45 % v Česku.

Regresná analýza

Ako bolo spomenuté, vstupné dáta IMD a IR nemajú normálne rozdelenie, a preto regresnou analýzou získané hodnoty regresných koeficientov majú len obmedzenú výpovednú hodnotu. Napriek tomu sú štandardnou súčasťou väčšiny doterajších hodnotení IMD (napr. Smith 2020, p. 50).

Aby bolo možné jasnejšie zachytiť zmeny medzi aktualizáčnymi verziami, skúmaniu regresnou analýzou je v tomto príspevku podrobená tesnosť a charakter vzťahu medzi referenčnými hodnotami nepriepustnosti IR_i a príslušnými hodnotami rozdielov, resp. chýb ($IMD_i - IR_i$). Ako vyplýva z parametrov regresnej (a súvisiacej korelačnej) analýzy v tab. 4, pri všetkých najnovších verziách v porovnaní s príslušnými predchádzajúcimi verziami je tento vzťah tesnejší (na Slovensku výrazne tesnejší), má opačný smer (pri najnovšej verzii negatívny, pri predchádzajúcich pozitívny), pričom sklon (b_1) poklesu regresnej priamky je strmší ako bol predtým sklon jej rastu (s výnimkou IR_A v Česku) a jej intercept (b_0) sa nachádza vyššie ako predtým (s výnimkou IR_A v Česku). Informácie o tomto aspekte tema-

tickej presnosti jednotlivých verzií vrstvy hustoty nepriepustnosti IMD sú pri jej aplikáciách veľmi dôležité a pri takej výraznej diskontinuite v jej vývoji, akú tieto dáta naznačujú, je potrebné na ňu pri prechode od starších verzií k novej upozorniť a v príslušných aplikáciách ju adekvátne reflektovať.

Tab. 4. Parametre regresnej analýzy vzťahu medzi referenčnými hodnotami nepriepustnosti IR_i a príslušnými hodnotami chýb, resp. rozdielov ($IMD_i - IR_i$)

| | | Slovensko | | | | | | Česko | | | | | |
|----------|----------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| | | (IMD 2006_v1) - (IR 2006_A/B) | (IMD 2006_v2) - (IR 2006_A/B) | (IMD 2006_v3) - (IR 2006_A/B) | (IMD 2006_v5) - (IR 2006_A/B) | (IMD 2009_v1) - (IR 2009_A/B) | (IMD 2009_v3) - (IR 2009_A/B) | (IMD 2006_v1) - (IR 2006_A/B) | (IMD 2006_v2) - (IR 2006_A/B) | (IMD 2006_v3) - (IR 2006_A/B) | (IMD 2006_v5) - (IR 2006_A/B) | (IMD 2009_v1) - (IR 2009_A/B) | (IMD 2009_v3) - (IR 2009_A/B) |
| IR2006_A | r | 0,330 | 0,349 | 0,369 | -0,637 | 0,387 | -0,665 | 0,140 | 0,237 | 0,242 | -0,317 | 0,272 | -0,324 |
| | r ² | 0,109 | 0,122 | 0,136 | 0,406 | 0,150 | 0,443 | 0,020 | 0,056 | 0,059 | 0,100 | 0,074 | 0,105 |
| | b ₁ | 0,230 | 0,225 | 0,227 | -0,262 | 0,238 | -0,270 | 0,108 | 0,138 | 0,137 | -0,136 | 0,151 | -0,136 |
| | b ₀ | -0,300 | -0,367 | -0,343 | -0,235 | -0,303 | -0,244 | -0,090 | -0,472 | -0,469 | -0,473 | -0,533 | -0,535 |
| IR2006_B | r | 0,255 | 0,259 | 0,277 | -0,692 | 0,337 | -0,693 | 0,113 | 0,196 | 0,200 | -0,359 | 0,228 | -0,371 |
| | r ² | 0,065 | 0,067 | 0,077 | 0,479 | 0,114 | 0,480 | 0,013 | 0,038 | 0,040 | 0,129 | 0,052 | 0,138 |
| | b ₁ | 0,168 | 0,160 | 0,162 | -0,301 | 0,202 | -0,294 | 0,085 | 0,112 | 0,111 | -0,155 | 0,124 | -0,157 |
| | b ₀ | -0,485 | -0,540 | -0,519 | -0,340 | -0,463 | -0,333 | -0,265 | -0,644 | -0,641 | -0,605 | -0,790 | -0,726 |

Zdroj: autori.

DISKUSIA A ZÁVER

Ak by sme za nepriepustné povrchy (ktoré by IMD mala zachytávať) nepovažovali aj nepriepustné povrchy s autochtóнным materiálom, teda „znepriepustnené“ vďaka aktivite človeka, napr. poľné cesty (z tohto predpokladu vychádza referenčná databáza IR_A), potom by tematická presnosť IMD bola vyššia v porovnaní s alternatívou, keď by sme nepriepustné povrchy s autochtóнным materiálom za nepriepustné povrchy (ktoré by IMD mala zachytávať) považovali (z tohto predpokladu vychádza referenčná databáza IR_B).

V oboch krajinách v predposledných (IMD2006_v3 a IMD2009_v1) a starších verziách sa početnosť pixelov malých hodnôt nepriepustnosti zdá byť podhodnotená a naopak početnosť pixelov veľkých hodnôt nepriepustnosti nadhodnotená. Toto nadhodnotenie je následne v najnovších verziách viac-menej primerane (v Česku), resp. až príliš (na Slovensku) redukované.

Na základe použitých hodnotiacich nástrojov sa zdá byť rozdiel v tematickej presnosti IMD v Česku a na Slovensku relatívne malý. V prípade najnovších verzií IMD2006_v5 a IMD2009_v3 (berúc do úvahy už len referenčnú databázu IR_A) je tematická presnosť IMD na Slovensku väčšia ako v Česku tak podľa korelačných

koeficientov (len s dvoma už spomenutými výnimkami), ako aj podľa veľkostí celkových absolútnych chýb relativizovaných, podľa veľkostí závažnejších chýb podhodnotenia a nadhodnotenia a tiež podľa ich podielu na celkovej rozlohe nesprávne klasifikovaných plôch. V oboch krajinách tematická presnosť IMD postupne narastá od starších verzií k novším – od IMD2006_v1 postupne až po IMD2006_v5 a od IMD2009_v1 po IMD2009_v3 – podľa všetkých týchto vymenovaných metód (s výnimkou ostatného menovaného podielu). Presnosť najnovšej verzie IMD pre rok 2006 (IMD2006_v5) sa zdá byť o málo menšia v porovnaní s presnosťou staršej dostupnej verzie IMD (IMD2009_v1) pre rok 2009 len v Česku a len pri použití Pearsonovho korelačného koeficientu. Podľa všetkých ostatných vymenovaných hodnotiacich nástrojov (s výnimkou ostatného menovaného podielu) je to v oboch krajinách naopak.

V súvislosti so záverom, že tematická presnosť jednotlivých verzií v čase narastá a výrazne sa zväčšila najmä pri porovnaní predposledných a príslušných najnovších aktualizovaných verzií, je potrebné upozorniť na dve skutočnosti, ktoré tento na prvý pohľad jasný a pozitívny záver do istej miery narúšajú. Prvou z nich je zväčšenie rozdielu medzi najpresnejším dostupným odhadom rozlohy nepriepustných povrchov (asi 1 143 km² v roku 2006 a 1 179 km² v roku 2009 na Slovensku a asi 2 532 km² v roku 2006 a 2 618 km² v roku 2009 v Česku, tab. 3) vychádzajúcim z priemeru vzorky príslušných referenčných databáz IR_A (2,33 % v roku 2006 a 2,40 % v roku 2009 na Slovensku a 3,21 % v roku 2006 a 3,32 % v roku 2009 v Česku – obr. 2) na jednej strane a odhadom rozlohy nepriepustných povrchov z priemeru vzorky príslušných validovaných databáz IMD (obr. 2) na strane druhej práve v súvislosti so zverejnením najnovších hodnotených verzií. Napríklad podľa priemeru vzorky staršej verzie IMD2009_v1 by mala byť rozloha nepriepustných povrchov na Slovensku (2,67 %) a v Česku (3,29 %) spolu asi 3 904 km², no podľa priemeru vzorky IMD2009_v3 by to malo byť na Slovensku (1,51 %) a v Česku (2,33 %) spolu len asi 2 578 km², čo je výrazne menej ako najpresnejší dostupný odhad podľa priemeru vzorky IR2009_A 3 795 km² (príp. IR2006_B 4 143 km²). Z ostatných štyroch percentuálnych hodnôt spadá len hodnota pre IMD2009_v1 v Česku (3,29 %) do príslušného 95-percentného intervalu spoľahlivosti odhadu priemeru vzorky IR2009_A (3,17 %, 3,47 %). Pre porovnanie Gallego et al. (2016, p. 6) uvádzajú tri rôzne odhady rozlohy nepriepustných povrchov pre rok 2012 v týchto dvoch krajinách spolu: 3 989 km², 5 410 km² a 5 350 km². V tomto smere sa teda zdá, že najnovšia verzia IMD je od reality vzdialenejšia ako tie predchádzajúce. Druhou skutočnosťou je výrazná zmena tesnosti a charakteru vzťahu medzi referenčnými hodnotami nepriepustnosti IR_i a príslušnými hodnotami rozdielov, resp. chýb ($IMD_i - IR_i$), ktorá opäť prišla práve v súvislosti so zverejnením najnovších verzií.

Ako už bolo naznačené, z pohľadu veľkosti celkovej absolútnej chyby relativizovanej TAER tematická presnosť aktualizovaných verzií v čase narastá. V najnovších verziách IMD2006 a IMD2009 na Slovensku a v Česku je celková absolútna chyba TAER, t. j. pomer rozlohy nesprávne klasifikovaných plôch a rozlohy nepriepustných povrchov v danej krajine, relativizovaná a menšia ako v tých predchádzajúcich – asi 44 až 46 %. Z toho asi dve pätiny alebo viac tvoria tzv. závažnejšie chyby (z pohľadu potenciálneho využitia IMD pre dezagregáciu sociálno-ekonomických premenných) vyskytujúce sa tak v štvorcoch 100 × 100 m bez nepriepustných povrchov, v ktorých IMD indikuje nenulovú nepriepustnosť (tvoriacich asi 1 až 2 % nesprávne klasifikovaných plôch), ako aj v štvorcoch 100 × 100 m

s istým, zväčša malým nenulovým podielom nepriepustných povrchov, v ktorých IMD indikuje nulovú nepriepustnosť (tvoriacich asi 38 až 45 % nesprávne klasifikovaných plôch).

Gallego et al. (2016, p. 3) odhadujú pre územie IMD pokrývajúce väčšinu územia Európy chybu vynechania 40,2 % a chybu priradenia 21,9 %. Hoci neudávajú, či to platí pre IMD2006, IMD2009 alebo IMD2012, podľa roku vydania ich štúdie je zrejmé, že vychádzajú z aktualizáčnych verzií IMD2006_v4, IMD2009_v2 alebo IMD2012_v1 publikovaných v roku 2014, ktoré žiaľ EEA odmietla poskytnúť na porovnanie s ostatnými verziami. Keďže nimi chápaná chyba vynechania je to isté ako nami definovaná celková absolútna chyba relativizovaná pre chyby podhodnotenia TAERU, pre porovnanie uvádzame príslušné hodnoty pre predchádzajúcu a nasledujúcu aktualizáčnú verziu tak pre rok 2006, ako aj pre rok 2009 v poradí IMD2006_v3, IMD2006_v5, IMD2009_v1, IMD2009_v3 (v %): Slovensko 26,3, 40,8, 24,3, 40,8; Česko 28,8, 37,3, 28,6, 37,8. Keďže z nami definovanej celkovej absolútnej chyby relativizovanej pre chyby nadhodnotenia (TAERO) je možné jednoduchým už popísaným spôsobom vypočítať obdobný ukazovateľ, a to Gallegom et al. (2013, p. 3) chápanú chybu priradenia, pre porovnanie uvádzame aj jej hodnoty v rovnakom poradí (%): Slovensko 31,7, 7,1, 32,0, 5,7; Česko 28,1, 12,5, 28,0, 11,5. Najnovšie aktualizáčné verzie teda vychádzajú z pohľadu Gallegom et al. (2016, p. 3) chápanej chyby priradenia najlepšie, no z pohľadu nimi chápanej chyby vynechania najhoršie.

Z uvedeného vyplýva, že v najnovších aktualizáčnych verziiach IMD pre roky 2006 a 2009 (publikovaných v roku 2019 pri príležitosti generovania IMD2015) sa v porovnaní s predchádzajúcimi verziami síce podarilo veľmi výrazne zmenšiť chyby nadhodnotenia, no na úkor výrazného zväčšenia chýb podhodnotenia. Prevalu chýb podhodnotenia v týchto verziiach naznačujú aj príslušné záporné smernice regresných priamok (striedajúce pozitívne smernice v predchádzajúcich aktualizáčnych verziiach) v tab. 4, z ktorých sa zdá, že s rastúcou intenzitou nepriepustnosti v referenčných dátach sa zväčšovala aj miera podhodnotenia IMD2006 a IMD2009 v porovnaní s príslušnými referenčnými dátami (pričom v predchádzajúcich aktualizáčnych verziiach sa s rastúcou intenzitou nepriepustnosti v referenčných dátach zväčšovala skôr miera nadhodnotenia).

Z hodnôt 0,41 až 0,89 smerníc regresie medzi IMD2015, resp. IMD2018 a príslušnými referenčnými databázami vytvorenými pomocou LUCAS v jednotlivých krajinách (resp. regiónoch) pokrytia IMD vyplýva, že stav prevažujúceho podhodnotenia pretrváva aj v IMD2015 a IMD2018 (Smith 2020, p. 50). Pozitívne je, že medzi týmito dvoma rokmi možno pozorovať opačný trend, hoci len v oveľa menšej miere: mierny pokles chyby vynechania na úkor mierneho rastu chyby priradenia (ibid. p. 2). Zrejmý je aj z nárastu hodnôt zo spomenutých smerníc z 0,41 až 0,79 pre rok 2015 na 0,56 až 0,89 pre rok 2018 (ibid. p. 49-50).

Ak by v záujmovom území s nenulovou výmerou nepriepustných povrchov x km² vrstva IMD indikovala nulovú výmeru nepriepustných povrchov, hodnota príslušnej celkovej absolútnej chyby relativizovanej TAER by bola 100 %, z čoho by závažnejšie chyby podhodnotenia (MaU) tvorili 100 %. Ak by v tom istom území iná verzia vrstvy IMD síce indikovala správnu celkovú výmeru nepriepustných povrchov x km² (tvoriacu maximálne polovicu rozlohy záujmového územia), avšak nepriepustné povrchy by indikovala len v tých častiach záujmového územia, kde sa v skutočnosti žiadne nepriepustné povrchy nenachádzajú, potom by hodnota príslušnej TAER bola 200 %, z čoho by jednu polovicu tvorili závažnejšie chyby pod-

hodnotenia (MaU) a druhú polovicu závažnejšie chyby nadhodnotenia (MaO). Keďže bolo preukázané, že hodnoty týchto charakteristík tematickej presnosti pre Slovensko a Česko pre roky 2006 a 2009 dosahujú výrazne nižšie úrovne, je možné vrstvu nepriepustnosti IMD pre dané krajiny a dané roky považovať za zdroj informácií s pridanou hodnotou, ktorý si nájde široké uplatnenie, avšak mal by byť používaný s rozvahou a obozretnosťou, s ohľadom na danú aplikáciu, požadovanú presnosť a mierku, resp. podrobnosť priestorového rozlíšenia (Hurbánek et al. 2010).

Kvalita výstupov v minulosti realizovaných aplikácií využívajúcich IMD2006 a IMD2009 na Slovensku a v Česku ako vstupy môže byť spätne prehodnotená aj na základe poznania konkrétnej aktualizáčnej verzie IMD, ktorá do týchto aplikácií vstupovala, keďže jednotlivé verzie – ako sa ukázalo – sa v rôznych charakteristikách tematickej presnosti do značnej miery líšia.

Tento článok vznikol aj vďaka podpore Agentúry na podporu výskumu a vývoja pre projekt č. APVV-20-0586.

LITERATÚRA

- BATISTA E SILVA, F., GALLEGU, J., LAVALLE, C. (2013). A high-resolution population grid map for Europe. *Journal of Maps*, 9, 16-28. DOI: <https://doi.org/10.1080/17445647.2013.764830>
- CONGEDO, L., SALLUSTIO, L., MUNAFÒ, M., OTTAVIANO, M., TONTI, D., MARCHETTI, M. (2016). Copernicus high-resolution layers for land cover classification in Italy. *Journal of Maps*, 12, 1195-1205. DOI: <https://doi.org/10.1080/17445647.2016.1145151>
- ČÚŽK (2022a). *Archivní ortofoto České republiky*, [Online]. Dostupné na: [https://geoportal.cuzk.cz/\(S\(c1a4mo45343dp4ruv1bg1m55\)\)/Default.aspx?lng=CZ&mode=TextMeta&side=ortofoto&metadataID=CZ-CUZK-ORTOARCHIV-R&mapid=83&menu=233](https://geoportal.cuzk.cz/(S(c1a4mo45343dp4ruv1bg1m55))/Default.aspx?lng=CZ&mode=TextMeta&side=ortofoto&metadataID=CZ-CUZK-ORTOARCHIV-R&mapid=83&menu=233) [cit: 3-1-2022].
- ČÚŽK (2022b). *Prohlížeč služba WMS – Archivní Ortofot*, [Online]. Dostupné na: https://geoportal.cuzk.cz/WMS_ORTOFOTO_ARCHIV/WMSservice.aspx?service=WMS&request=GetCapabilities [cit: 3-1-2022].
- EEA (2006). *GMES fast track service precursor on land monitoring. EEA proposal 2006 – 2008 – Terms of reference*. Copenhagen (EEA), p. 10 [Online]. Dostupné na: http://www.gmes-geoland.info/download/ToR_FTSP-LandMonitoring.pdf [cit: 19-12-2020].
- EEA (2007). *Tender specifications – GMES fast track service precursor on land monitoring (EEA/IDS/07/001)*. Copenhagen (EEA), p. 16, [Online]. Dostupné na: <http://www.eea.europa.eu/about-us/tenders/EEAIDS07001> [cit: 19-12-2020].
- EEA (2009). *Tender specification for SSL – GMES FTS precursor on land monitoring (EEA/SES/09/003)*. Copenhagen (EEA), [Online]. Dostupné na: <http://www.eea.europa.eu/about-us/tenders/eea-ses-09-003/tender-specifications-eea-ses-09.003/view> [cit: 19-12-2020].
- EEA (2017). *INDICATOR ASSESSMENT – Imperviousness and imperviousness change* Copenhagen (EEA), [Online]. Dostupné na: <https://www.eea.europa.eu/data-and-maps/indicators/imperviousness-change-1/assessment> [cit: 19-12-2020].
- EEA (2020). *Copernicus land monitoring service – High resolution layers – Imperviousness*, [Online]. Dostupné na: <https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-imperviousness-2> [cit: 20-12-2020].
- EUROSENSE a GEODIS SLOVAKIA (2007). *Súvislá digitálna ortofotomapa SR z r. 2005 – 2007 v rozlíšení 50 cm*, [Online]. Dostupné na: <https://www.eurosense.sk/kopie-ortofotomapy-2d-3d-dtm> [cit: 13-2-2018].

- EUROSENSE a GEODIS SLOVAKIA (2010). *Súvislá digitálna ortofotomapa SR z r. 2008 – 2010 v rozlíšení 25 cm*, [Online]. Dostupné na: <https://www.eurosense.sk/kopie-ortofotomapy-2d-3d-dtm> [cit: 11-3-2018].
- FAO (2005). *Land cover classification system. Classification concepts and user manual*. Rome (FAO and UN Rome), [Online]. Dostupné na: <https://wedocs.unep.org/handle/20.500.11822/29692> [cit: 19-12-2020].
- GALLEGO, J., SANNIER, CH., PENNEC, A., DUFOURMONT, H. (2016). *Validation of Copernicus land monitoring services and area estimation*, [Online]. Rome: ICASVII, 2016. Dostupné na: <https://land.copernicus.eu/user-corner/technical-library/validation-of-copernicus-land-monitoring-services-and-area-estimation/view> [cit: 19-12-2020].
- GANGKOFNER, U., WEICHELBAUM, J., KUNTZ, S., BRODSKY, L., LARSSON, K., De PASQUALE, V. (2010). Update of the European high-resolution layer of built-up areas and soil sealing 2006 with image 2009 data. *30th EARSeL Symposium 2010: Remote Sensing for Science, Education and Culture, 31 May – 3 June 2010. Paris, France*.
- GSE LAND (2008). Brochure InfoSheet impervious areas and sealing levels. In *GSE Land Informations Services website, Communications, Promotion material*. GSE Land project, 2008, p. 2.
- HURBÁNEK, P., ATKINSON, P. M., PAZÚR, R., ROSINA, K., CHOCKALINGAM, J. (2010). Accuracy of built up area mapping in Europe at varying scales and thresholds. In Tate, N. J., Fisher, P. F., eds. *Accuracy 2010: Proceedings of the ninth International symposium on spatial accuracy assessment in natural resources and environmental sciences, 20-23 July 2010*. Leicester (University of Leicester, International Spatial Accuracy Research Association – ISARA), pp. 385-388.
- HURBÁNEK, P., ROSINA, K. (2013). Metódy pre hodnotenie tematickej presnosti Európskej mapy umelých nepriepustných povrchov. In *Zborník abstraktov – Exkurzný sprievodca*. Ružomberok (VERBUM – vydavateľstvo KU), pp. 49-52.
- KOLÁŘ, J. 2003. *Geografické informační systémy*. Praha (ČVUT).
- KOPECKÝ, M., KAHABKA, H. (2008). *Delivery report Slovakia: Issue 1.0*. Potsdam (InfoTerra GmbH).
- KRÓWCZYŃSKA, M., SOSZYŃSKA, A., PABJANEK, P., WILK, E., HURBÁNEK, P., ROSINA, K. (2016). Accuracy of the soil sealing enhancement product for Poland. *Quaestiones Geographicae*, 35, 89-95, DOI: <https://doi.org/10.1515/quageo-2016-0029>
- LANGANKE, T., STEIDL, M., SCHLEICHER, CH., SANNIER, CH. (2016). *Copernicus land monitoring service – High resolution layer imperviousness: Product specifications document*, [Online]. Dostupné na: <https://land.copernicus.eu/user-corner/technical-library/hrl-imperviousness-technical-document-prod-2015> [cit: 19-12-2020].
- LEFEBVRE, A., BEAUGENDRE, N., PENNEC, A., SANNIER, C., CORPETTI, T., (2013) Using data fusion to update built-up areas of the 2012 European high-resolution layer imperviousness. In Lasaponara, R., Masini, N., Biscione, M., eds. *Proceedings of the 33th EARSeL Symposium Towards Horizon 2020: Earth Observation and Social Perspectives, Matera, Italy, 3 – 6 June 2013*. Matera (European Association of Remote Sensing Laboratories), pp. 321-328.
- LEHNER, A., NAEIMI, V., STEINNOCHER, K. (2017). Sentinel-1 for urban areas – Comparison between automatically derived settlement layers from Sentinel-1 Data and Copernicus High Resolution Information Layers. In Ragia, L., Rocha, J. G., Laurini, R., eds. *Proceedings of the 3rd International Conference on Geographical Information Systems Theory, Applications and Management – GISTAM – Volume 1*. Setúbal (SciTePress), pp. 43-49. DOI: <https://doi.org/10.5220/0006320800430049>
- MACHKOVÁ, N., NOVÁČEK, J. (2009). CORINE – mapovanie Slovenska. *Enviromagazín*, 3, 12-13.
- MACHKOVÁ, N., NOVÁČEK, J., TOBIK, J. (2008). *Verification of high resolution soil sealing layer qualitative assessment prepared by Slovak Republic*. Banská Bystrica (Slovenská agentúra životného prostredia).

- MAUCHA, G., BÜTTNER, G. (2008). *Recommendations for quantitative assessment of high-resolution soil sealing layer, v. 2*. Copenhagen (EEA and European Topic Centre for Land Use and Spatial Information).
- MAUCHA, G., BÜTTNER, G., KOSZTRA, B. (2010). *European validation of GMES FTS Soil Sealing Enhancement Data (final draft)*. Budapest (FOMI).
- POHLOVÁ, R. (2013). *Hodnocení přesnosti databáze zastavěných území v Česku*. Bakalářská práce, Univerzita Karlova, Praha.
- PROJECT MANAGER COPERNICUS LAND SERVICES (2020). *Copernicus Land monitoring services: Imperviousness*, [Online]. Dostupné na: <https://land.copernicus.eu/pan-european/high-resolution-layers/imperviousness> [cit: 13-1-2021].
- ROSINA, K., HURBANEK, P. (2016). *Spatial disaggregation of population density using land cover and remote sensing data*. Geographia Slovaca, 31. Bratislava (Geografický ústav SAV).
- SANNIER, CH., GALLEGO, J., DAHMER, J., SMITH, G., DUFOURMONT, H., PENNEC, A. (2016). Validation of Copernicus High Resolution Layer on Imperviousness degree for 2006, 2009 and 2012. In Bailly, J. S., Griffith, D., Josselin, D., eds. *Proceedings of Spatial Accuracy 2016, 5 – 8 July 2016. Montpellier, France*. Montpellier (UMR 7300 and ESPACE), pp. 60-66.
- SMITH, G. (2020). *GMES Initial Operations / Copernicus Land monitoring services – Validation of products. HRL Imperviousness degree 2018 Validation report MDI/14/010*, [Online]. Copenhagen (EEA). Dostupné na: https://land.copernicus.eu/user-corner/technical-library/clms_hrl_imd_validation_report_sc04_1_3.pdf [cit: 13-1-2021].
- STEENMANS, Ch., SOUSA, A. (2007). *Guidelines for verification of high resolution soil sealing layer: Qualitative assessment*. Copenhagen (EEA).
- STEHMAN, S. V. (1999). Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, 20, 2423-2441.
- STEINNOCHER, K., KÖSTL., M., WEICHELBAUM, J. (2011). Grid-based population and land take trend indicators – New approaches introduced by the geoland2 Core Information Service for Spatial Planning. *New techniques and technologies for statistics NTS conference*. Brussels (Eurostat). (CD-ROM).
- XIAN, G., HOMER, C., DEWITZ, J., FRY, J., HOSSAIN, N., WICKHAM, J. (2011). Change of impervious surface area between 2001 and 2006 in the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, 77, 758-762.

Andrea Čorbová, Pavol Hurbánek, Konštantín Rosina

THEMATIC ACCURACY DEVELOPMENT OF IMPERVIOUSNESS DENSITY LAYER FOR YEARS 2006 AND 2009 IN SLOVAKIA AND CZECHIA

The demand for current, accurate and high-quality geodata is constantly growing, and therefore the European Union has stepped in with support of increased Earth observation geodata production and application. Some of the most used geodata come from the CORINE Land Cover (CLC) project, which has been operating since 1985 and its latest output was produced for the year 2018. The biggest disadvantages of this data layer are the minimum mapping unit of 25 ha and the update frequency of six years. Together with its update for the year 2006, the CLC source satellite imagery was also used to generate Imperviousness Density High Resolution Layer (IMD) – the raster with a full 20 × 20 m resolution as an unvalidated intermediate product and an aggregated 100 × 100 m resolution as a validated final product with pixel integer values from 0% to 100% representing the imperviousness degree. It was the first of the high resolution layers focused on a single land cover class later produced within the Copernicus Programme. As the knowledge of the spatial

distribution of impervious surfaces is indispensable for many human activities the IMD has a wide range of potential uses. Nowadays, it is available for the reference years 2006, 2009, 2012, 2015 and 2018.

The objective of the paper is to analyse various aspects of the thematic accuracy of 4 of the 5 successively published update versions of IMD2006 and 2 of the 3 successively published update versions of IMD2009:

IMD2006_v1 – available from June 12, 2009, so-called “precursor”,

IMD2006_v2 – available from January 26, 2010, so-called “enhancement”,

IMD2006_v3, IMD2009_v1 – available from January 22, 2013, published on the occasion of IMD2009 production,

IMD2006_v4, IMD2009_v2 – available around 2014, on the occasion of generating IMD2012 (We were unable to obtain these at the time of their availability. EEA have chosen not to make older versions of these two products available after their latest update versions are published.),

IMD2006_v5, IMD2009_v3 – available from January 15, 2019, published on the occasion of IMD2015 production.

It is obvious that the gradual development of these update versions leads to a continuing improvement of their thematic accuracy, but the extent and various aspects of this improvement have not been known until now. The thematic accuracy is analysed with an emphasis on the application of IMD in the spatial disaggregation of socioeconomic, most often population variables, for example as auxiliary data in dasymetric mapping. It was for this purpose that some of the update versions of IMD2006 and IMD2009 have been used in the past, e.g. by the Austrian Institute of Technology (Steinnocher et al. 2011), the Joint Research Center of the European Commission (Batista e Silva et al. 2013) and others (e.g. Rosina and Hurbánek 2016). For the professional community that uses the population grids produced within these (or other) studies (that have been or will be published), it is desirable to know the quality of the update versions employed as inputs in these studies. Unlike the IMD for the years 2012, 2015 and 2018, which have been validated several times (e.g. Congedo et al. 2016, Smith 2020 and Lehner et al. 2017), thanks to the significant increase in the availability of alternative data sources on the spatial distribution of impervious surfaces in the second decade of this millennium, the IMD for the years 2006 and 2009 lag behind in this regard. The main reason is probably the lack of alternative data sources for these two years, which could be used as reference, and thus the difficulty of such validation inhering in the need of preparation of a reference database for this specific purpose.

For almost a decade, most of the IMD validations have been based on stratified systematic samples of 100 m pixels using Land Use/Cover Area Frame Survey (LUCAS) with relatively small sample sizes, each pixel with 25 sampling points. In Slovakia and Czechia, which are the study areas in this paper, they relied on about several hundred sample pixels. However, this approach – similarly as the one based on a stratified random sample (Maucha and Büttner 2008) – does not provide a reliable estimate of omission error for a class covering only a small fraction of the total area (Maucha et al. 2010, p. 15, Sannier et al. 2016, p. 64-66 and Smith 2020, p. 18), which is important especially in applications, where the validated product is used as an ancillary dataset in population and other socioeconomic data disaggregation.

Therefore, a large simple random sample of 20,000 pixels with 100 m spatial resolution per country per year, each pixel with 100 sampling points, is analysed in this study employing histograms, correlation and regression analyses, relative total absolute error (TAER, i.e. the ratio of the area of incorrectly classified surfaces to the area of impervious surfaces in a given country and a given year), and its structure, namely the proportions of major underestimation TAER ($TAER_{MaU}$ for $IMD_i - IR_i$ differences when $IMD_i = 0$ and $IR_i > 0$), major overestimation TAER ($TAER_{MaO}$ for $IMD_i - IR_i$ differences when $IMD_i > 0$ and $IR_i = 0$), minor underestimation TAER ($TAER_{MIU}$ for negative $IMD_i - IR_i$ differences when $IMD_i > 0$)

and $IR_i > 0$), and minor overestimation TAER ($TAER_{MiO}$ for positive $IMD_i - IR_i$ differences when $IMD_i > 0$ and $IR_i > 0$) to TAER. Please note that the term “major” corresponds to the viewpoint of the potential application of the IMD in socioeconomic data disaggregation.

In both countries, in the penultimate (IMD2006_v3 and IMD2009_v1) and older versions, the number of pixels with small imperviousness values seems to be underestimated, and conversely, the number of pixels with large imperviousness values appears to be overestimated. While in the latest versions this overestimation seems to be reduced just about right in Czechia, it appears to be reduced too much in Slovakia.

Overall, the IMD thematic accuracy in Slovakia and Czechia seems to be similar. In the latest versions IMD2006_v5 and IMD2009_v3 though, the accuracy in Slovakia appears somewhat better according to most of the correlation coefficient values, TAER, $TAER_{MaU}$ and $TAER_{MaO}$ values, and also the proportion value of $TAER_{Ma}$ ($TAER_{MaU} + TAER_{MaO}$) to TAER. In both countries, the IMD thematic accuracy gradually increases with almost each new published version (especially with the latest one) of the product according to the criteria listed, except for the proportion of $TAER_{Ma}$ to TAER.

However, the difference between the most accurate available estimate of the area of impervious surfaces (about 1,143 km² in 2006 and 1,179 km² in 2009 in Slovakia and about 2,532 km² in 2006 and 2,618 km² in 2009 in Czechia – Tab. 3) calculated from the IR_A reference database sample average (2.3% in 2006 and 2.4% in 2009 in Slovakia and 3.2% in 2006 and 3.32% in 2009 in Czechia – Fig. 2) on the one hand, and the estimate calculated from the corresponding IMD validated database sample average (Fig. 2) on the other hand, is much greater in the latest update versions than in the previous ones. For example, according to the older version IMD2009_v1 the joint area of impervious surfaces in these two countries amounts to about 3,904 km² (2.7% Slovakia and 3.3% Czechia), but according to the latest version IMD2009_v3 it amounts to about 2,578 km² (1.5% in Slovakia and 2.3% in Czechia), which is notably less than the most accurate available estimate according to the IR2009_A reference database 3,795 km². Of the latter four percentage values, only the value for IMD2009_v1 in Czechia (3.3%) falls within the corresponding 95% confidence interval of the IR2009_A sample mean estimate of 3.3% (3.2% and 3.5%). For comparison, Gallego et al. (2016, p. 6) give three different estimates of the joint area of impervious surfaces in the two countries for the year 2012: 3,989 km², 5,410 km² and 5,350 km². In this respect, it seems that the latest version of IMD is further from reality than the previous ones.

TAER values for Slovakia and Czechia in the latest IMD2006 and IMD2009 update versions (about 44 to 46%) are smaller than in the corresponding previous versions. Of this, $TAER_{Ma}$ forms about two-fifths or more. Misclassified surfaces corresponding to $TAER_{MaO}$ occur in pixels without impervious surfaces, in which the IMD indicates non-zero imperviousness, and they account for about 1 to 2% of the total area of misclassified surfaces. Misclassified surfaces corresponding to $TAER_{MaU}$ occur in pixels with a certain, mostly small, non-zero proportion of impervious surfaces, in which the IMD indicates zero imperviousness, and they account for about 38 to 45% of the total area of misclassified surfaces.

Gallego et al. (2016, p. 3) estimate the omission error of 40.2% and the commission error of 21.9% for the IMD coverage of most of Europe. According to the publication year of their study, it is clear these error values apply to update version IMD2006_v4, IMD2009_v2 or IMD2012_v1 published in 2014, which have not been provided by the EEA for comparative analysis presented here. Since the definition of the omission error by Gallego et al. (2016, p. 3) is the same as the definition of $TAER_U$ ($TAER_{MaU} + TAER_{MiU}$) here, for comparison, we list the TAERU values for the previous and next update versions for both 2006 and 2009 in the following order – IMD2006_v3, IMD2006_v5, IMD2009_v1, IMD2009_v3 (in %): Slovakia 26.3, 40.8, 24.3, 40.8; Czechia 28.8, 37.3, 28.6, 37.8. Since $TAER_O$ ($TAER_{MaO} + TAER_{MiO}$) presented here can be easily recalculated to the commission error defined by Gallego et al. (2016, p. 3), for comparison, we also list these commission error values in the same order (in %): Slovakia 31.7, 7.1, 32.0, 5.7; Czechia

28.1, 12.5, 28.0, 11.5. Thus, the latest IMD2006 and IMD2009 update versions – when compared to the previous ones in these two countries – score the best with respect to the commission error, but the worst with respect to the omission error (when both are defined as in the aforementioned study). In other words, these latest versions have brought about a considerable reduction in overestimation errors at the expense of a considerable increase in underestimation errors.

Similar indications come from the correlation and regression analyses of the relationship between the IR_i reference values and the respective $IMD_i - IR_i$ difference values (Tab. 4), which suggest that for all the latest IMD2006 and IMD2009 update versions in these two countries this relationship is tighter compared to the corresponding previous versions and it has the opposite direction than before: negative for the latest version, positive for the previous ones.

The regression slope values of 0.41 to 0.89 between IMD2015 and IMD2018 on the one hand and the corresponding reference data based on LUCAS on the other hand in individual countries (or regions) of the IMD coverage presented by Smith (2020, p. 50) indicate that the predominance of underestimation persists in the IMD even in these two years. On a positive note, the opposite trend can be observed between these two years, although only to a much lesser extent – a slight decrease in the omission error at the expense of a slight increase in the commission error (ibid. p. 2) – as the regression slope values grew from about 0.41 to 0.79 in 2015 to about 0.56 to 0.89 in 2018 (ibid. p. 49-50).

If the IMD indicated zero impervious surface area in a hypothetical study area with a non-zero total impervious surface area of $x \text{ km}^2$, then the corresponding TAER value would be 100%, of which the $TAER_{MaU}$ proportion would be 100%. If, in the same study area, another IMD version indicated the correct total impervious surface area of $x \text{ km}^2$ (where x is not more than half of the total area of the study area), but it would indicate impervious surfaces only in those parts of the study area where no impervious surfaces actually exist, then the corresponding TAER value would be 200%, of which the $TAER_{MaU}$ and $TAER_{MaO}$ proportions would be one half each. Because these characteristics of the IMD thematic accuracy for Slovakia and Czechia for the years 2006 and 2009 reach significantly lower levels, the IMD for the given countries and given years can be considered a data source with added value that finds wide application, but should be used with caution and care, taking into account the given application, its scale (Hurbánek et al. 2010) and the required accuracy.



Article first received: May 2022
Article accepted: June 2023